

Thu Oct 14 09:29:46 2004

us-10-021-002-2.ra1

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 18:08:19 : Search time 21 Seconds  
(without alignments)  
997.928 Million cell updates/sec

Title: US-10-021-002-2  
Perfect score: 1659  
Sequence: 1 MVELMPFLILLPFLYMA.....RNETIARLMDVITCWSAQ 316

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pdp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pdp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pdp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pdp.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pdp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfilest.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1659	100.0	316 1 US-08-464-400-2	Sequence 2, Appli
2	1659	100.0	316 3 US-08-875-273A-2	Sequence 2, Appli
3	1659	100.0	316 3 US-09-123-366-2	Sequence 2, Appli
4	1659	100.0	316 5 PCT-US95-0187A-2	Sequence 2, Appli
5	1577	95.1	318 3 US-09-439-313-339	Sequence 338, App
6	1577	95.1	318 3 US-09-352-616A-339	Sequence 338, App
7	1577	95.1	318 4 US-09-636-215-339	Sequence 339, App
8	1577	95.1	318 4 US-09-685-166A-339	Sequence 339, App
9	1577	95.1	318 4 US-09-679-426-339	Sequence 339, App
10	512.5	29.8	309 4 US-09-270-767-4172A	Sequence 4172A, A
11	494.5	29.8	323 2 US-08-580-545B-2	Sequence 2, Appli
12	494.5	29.8	323 3 US-09-462-655A-2	Sequence 2, Appli
13	474	28.6	330 4 US-09-489-847-162	Sequence 162, App
14	466.5	28.1	320 2 US-08-530-165-7	Sequence 7, Appli
15	454	27.4	330 4 US-09-489-847-330	Sequence 330, App
16	443	26.7	380 4 US-09-252-991A-32854	Sequence 32854, A
17	427.5	25.8	407 4 US-09-270-767-43967	Sequence 43967, A
18	381.5	23.0	274 4 US-09-328-352-8213	Sequence 8213, Ap
19	359.5	21.7	287 2 US-08-901-306-2	Sequence 2, Appli
20	354.5	21.4	287 3 US-09-180-271-2	Sequence 2, Appli
21	275.5	16.6	336 4 US-09-248-796A-15559	Sequence 15559, A
22	262.5	15.8	354 4 US-09-248-796A-16822	Sequence 16822, A
23	260	15.7	401 4 US-09-248-796A-16145	Sequence 16145, A
24	254	15.3	503 4 US-09-248-796A-15340	Sequence 15340, A
25	254	15.3	507 4 US-09-792-024-109	Sequence 109, App
26	238	14.3	148 3 US-09-724-864-41	Sequence 41, Appli
27	229	13.8	203 4 US-09-634-955B-26	Sequence 26, Appli

28	227	13.7	203 4 US-09-634-955B-17	Sequence 17, Appli
29	224.5	13.5	107 4 US-09-270-767-33908	Sequence 33908, A
30	224.5	13.5	107 4 US-09-270-767-49125	Sequence 49125, A
31	222	13.4	203 4 US-09-634-955B-20	Sequence 20, Appli
32	219	13.2	294 4 US-09-328-352-4992	Sequence 4992, Ap
33	215.5	13.0	238 4 US-08-586-664-1	Sequence 1, Appli
34	215.5	13.0	248 3 US-09-388-028-11	Sequence 11, Appli
35	215.5	13.0	248 4 US-09-726-614-11	Sequence 11, Appli
36	215.5	13.0	248 4 US-09-385-040-11	Sequence 3280, A
37	209	12.6	267 4 US-09-252-991A-23096	Sequence 23096, A
38	209	12.6	339 4 US-09-148-545-248	Sequence 248, App
39	206.5	12.4	632 4 US-09-328-352-6485	Sequence 6485, Ap
40	203.5	12.3	288 4 US-09-148-545-192	Sequence 192, App
41	203.5	12.3	318 4 US-09-148-545-192	Sequence 192, App
42	200	12.1	300 4 US-09-907-794A-159	Sequence 37, Appli
43	200	12.1	300 4 US-09-866-028-37	Sequence 159, App
44	200	12.1	300 4 US-09-905-125A-159	Sequence 159, App
45	200	12.1	300 4 US-09-902-775A-159	Sequence 159, App

# ALIGNMENTS

RESULT 1  
US-08-464-400-2  
Sequence 2, Application US/08464400  
Patent No. 5786204  
GENERAL INFORMATION:  
APPLICANT: HE, ET AL.  
TITLE OF INVENTION: Human Prostatic Specific Reductase  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
STREET: 6 BECKER FARM ROAD  
ADDRESS: CECCHI, STEWART & OLSTEIN  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,400  
FILING DATE: Concurrently  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-228  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1744  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 316 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN

Query Match 100.0%; Score 1659; DB 1; Length 316;  
Best local Similarity 100.0%; Pred. No. 1.1e-186; Index 0; Gaps 0;  
Matches 316; Conservative 0; Mismatches 0;

1 MVELMPFLILLPFLYMAPOIRKMLSGVCTSTVQLPGKVVTGANTGIGKETAKE 60  
|||||

Db 1 MWELMPFLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVVTGANTGIGKETAKE 60  
QY 61 LAORGARVYLACRDVEKSELVAKETIOTTGNOQVLRKLDSDTKSIPAMAKGFAEKH 120  
Db 61 LAORGARVYLACRDVEKSELVAKETIOTTGNOQVLRKLDSDTKSIPAMAKGFAEKH 120  
QY 121 LHWINNAGVMMCPYSKTADGFEHMGVNHJGHFLTHLLEKLESAPSRIVNSSLAH 180  
Db 121 LHWINNAGVMMCPYSKTADGFEHMGVNHJGHFLTHLLEKLESAPSRIVNSSLAH 180  
QY 181 HLGRIHFHNLQGEKRYNAGLAYCHSKLANILFTQELARLKGSGVTTYSVHPGTVOSELY 240  
Db 181 HLGRIHFHNLQGEKRYNAGLAYCHSKLANILFTQELARLKGSGVTTYSVHPGTVOSELY 240  
QY 241 RHSSFRMMWMLFSPFITPQOGAOTRLHCAITGELTILSGNHPSDCHVAVMSQAQARNET 300  
Db 241 RHSSFRMMWMLFSPFITPQOGAOTRLHCAITGELTILSGNHPSDCHVAVMSQAQARNET 300  
QY 301 IARRLMDVITVTCMASQ 316  
Db 301 IARRLMDVITVTCMASQ 316

## RESULT 2

US-08-875-273A-2  
; Sequence 2, Application US/08875273A

; Patent No. 6106829  
; GENERAL INFORMATION:  
; APPLICANT: HE, ET AL.  
; TITLE OF INVENTION: Human Prostatic Specific Reductase  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/875,273A  
; FILING DATE: Concurrently  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FERRARO, GREGORY D.  
; REGISTRATION NUMBER: 36,114  
; REFERENCE/DOCKET NUMBER: 325800-228  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 316 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; US-08-875-273A-2

Query Match 100.0%; Score 1659; DB 3; Length 316;  
Best Local Similarity 100.0%; Pred. No. 1,1e-186;  
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWELMPFLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVVTGANTGIGKETAKE 60  
Db 1 MWELMPFLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVVTGANTGIGKETAKE 60

Db 1 MWELMPFLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVVTGANTGIGKETAKE 60  
QY 61 LAORGARVYLACRDVEKSELVAKETIOTTGNOQVLRKLDSDTKSIPAMAKGFAEKH 120  
Db 61 LAORGARVYLACRDVEKSELVAKETIOTTGNOQVLRKLDSDTKSIPAMAKGFAEKH 120  
QY 121 LHWINNAGVMMCPYSKTADGFEHMGVNHJGHFLTHLLEKLESAPSRIVNSSLAH 180  
Db 121 LHWINNAGVMMCPYSKTADGFEHMGVNHJGHFLTHLLEKLESAPSRIVNSSLAH 180  
QY 181 HLGRIHFHNLQGEKRYNAGLAYCHSKLANILFTQELARLKGSGVTTYSVHPGTVOSELY 240  
Db 181 HLGRIHFHNLQGEKRYNAGLAYCHSKLANILFTQELARLKGSGVTTYSVHPGTVOSELY 240  
QY 241 RHSSFRMMWMLFSPFITPQOGAOTRLHCAITGELTILSGNHPSDCHVAVMSQAQARNET 300  
Db 241 RHSSFRMMWMLFSPFITPQOGAOTRLHCAITGELTILSGNHPSDCHVAVMSQAQARNET 300  
QY 301 IARRLMDVITVTCMASQ 316  
Db 301 IARRLMDVITVTCMASQ 316

## RESULT 3

US-09-123-386-2  
; Sequence 2, Application US/09123386

; Patent No. 6344198  
; GENERAL INFORMATION:  
; APPLICANT: HE, ET AL.  
; TITLE OF INVENTION: Human Prostatic Specific Reductase  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
; STREET: 9410 KEY WEST AVE  
; CITY: ROCKVILLE  
; STATE: MARYLAND  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/123,386  
; FILING DATE: JUL-27-1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/464,400  
; FILING DATE: JUN-05-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jonathan L. Klein  
; REGISTRATION NUMBER: 41,119  
; REFERENCE/DOCKET NUMBER: PF150D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 316 AMINO ACIDS  
; TYPE: AMINO ACID  
; STRANDEDNESS:  
; TOPOLOGY: LINEAR  
; MOLECULE TYPE: PROTEIN  
; US-09-123-386-2

Query Match 100.0%; Score 1659; DB 3; Length 316;  
Best Local Similarity 100.0%; Pred. No. 1,1e-186;  
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MWELMPFLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVVTGANTGIGKETAKE 60  
Db 1 MWELMPFLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVVTGANTGIGKETAKE 60

QY 61 LAORGARVYIACRDVEKELVAKETOTTGNQCVLRKLDLSTKSIIRAMAKGFAEEKH 120  
DB 61 LAORGARVYIACRDVEKELVAKETOTTGNQCVLRKLDLSTKSIIRAMAKGFAEEKH 120  
QY 121 LHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNVSSLAH 180  
DB 121 LHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNVSSLAH 180  
QY 181 HLGRIHFHNLQGEKFXNAGLAICHSKLANILFTQELARLKGSGVTTYSVHPGTVOSELY 240  
DB 181 HLGRIHFHNLQGEKFXNAGLAICHSKLANILFTQELARLKGSGVTTYSVHPGTVOSELY 240  
QY 241 RHSSFRMMWMLPFSFIKTPOQGAQTRHLCALTEGLEILSGNHFSDCHVAVWSAQARNET 300  
DB 241 RHSSFRMMWMLPFSFIKTPOQGAQTRHLCALTEGLEILSGNHFSDCHVAVWSAQARNET 300  
QY 301 IARRLMDVITVCWASQ 316  
DB 301 IARRLMDVITVCWASQ 316

RESULT 4  
PCT-US95-01827A-2

Sequence 2, Application PC/TUS9501827A  
GENERAL INFORMATION:  
APPLICANT: HE, ET AL.  
TITLE OF INVENTION: Human Prostatic Specific Reductase  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/01827A  
FILING DATE: Concurrently  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-228  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 316 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
PCT-US95-01827A-2

Query Match 100.0%; Score 1659; DB 5; Length 316;  
Best Local Similarity 100.0%; Pred. No. 1,1e-186;  
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVELMFPILLILLPFLLYMAAPDIRKMLSSGVCTSTVQJPGKVVVVTGANTGIGKETAKE 60  
DB 1 MVELMFPILLILLPFLLYMAAPDIRKMLSSGVCTSTVQJPGKVVVVTGANTGIGKETAKE 60

QY 61 LAORGARVYIACRDVEKELVAKETOTTGNQCVLRKLDLSTKSIIRAMAKGFAEEKH 120  
DB 61 LAORGARVYIACRDVEKELVAKETOTTGNQCVLRKLDLSTKSIIRAMAKGFAEEKH 120  
QY 121 LHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNVSSLAH 180  
DB 121 LHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNVSSLAH 180  
QY 181 HLGRIHFHNLQGEKFXNAGLAICHSKLANILFTQELARLKGSGVTTYSVHPGTVOSELY 240  
DB 181 HLGRIHFHNLQGEKFXNAGLAICHSKLANILFTQELARLKGSGVTTYSVHPGTVOSELY 240  
QY 241 RHSSFRMMWMLPFSFIKTPOQGAQTRHLCALTEGLEILSGNHFSDCHVAVWSAQARNET 300  
DB 241 RHSSFRMMWMLPFSFIKTPOQGAQTRHLCALTEGLEILSGNHFSDCHVAVWSAQARNET 300  
QY 301 IARRLMDVITVCWASQ 316  
DB 301 IARRLMDVITVCWASQ 316

RESULT 5  
US-09-439-313-339

Sequence 339, Application US/09439313  
Patent No. 6329505  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, David C.  
APPLICANT: Hacham, Jennifer L.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang Yugu  
APPLICANT: Reed, Steven G.  
APPLICANT: Kalos, Michael  
APPLICANT: Fanger, Gary  
APPLICANT: Retter, Mark  
APPLICANT: Solk, John  
APPLICANT: Day, Craig  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
FILE REFERENCE: 210121.42709  
CURRENT APPLICATION NUMBER: US/09/439,313  
CURRENT FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 575  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 339  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-439-313-339

Query Match 95.1%; Score 1577; DB 3; Length 318;  
Best Local Similarity 98.7%; Pred. No. 5,4e-177;  
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MVELMFPILLILLPFLLYMAAPDIRKMLSSGVCTSTVQJPGKVVVVTGANTGIGKETAKE 60  
DB 1 MVELMFPILLILLPFLLYMAAPDIRKMLSSGVCTSTVQJPGKVVVVTGANTGIGKETAKE 60  
QY 61 LAORGARVYIACRDVEKELVAKETOTTGNQCVLRKLDLSTKSIIRAMAKGFAEEKH 120  
DB 61 LAORGARVYIACRDVEKELVAKETOTTGNQCVLRKLDLSTKSIIRAMAKGFAEEKH 120  
QY 121 LHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNVSSLAH 180  
DB 121 LHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNVSSLAH 180  
QY 181 HLGRIHFHNLQGEKFXNAGLAICHSKLANILFTQELARLKGSGVTTYSVHPGTVOSELY 240  
DB 181 HLGRIHFHNLQGEKFXNAGLAICHSKLANILFTQELARLKGSGVTTYSVHPGTVOSELY 240  
QY 241 RHSSFRMMWMLPFSFIKTPOQGAQTRHLCALTEGLEILSGNHFSDCHVAVWSAQARNET 300  
DB 241 RHSSFRMMWMLPFSFIKTPOQGAQTRHLCALTEGLEILSGNHFSDCHVAVWSAQARNET 300

QY 301 IARRLMDV 308  
 Db 301 IARRLMDV 308

## RESULT 6

US-09-352-616A-339  
 ; Sequence 339, Application US/09352616A  
 ; Patent No. 6395278  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Harlocker, Susan Louise  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Xu, Jiangchun  
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
 ; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: 210121.427C8  
 ; CURRENT APPLICATION NUMBER: US/09/352,616A  
 ; CURRENT FILING DATE: 1999-07-13  
 ; NUMBER OF SEQ ID NOS: 472  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 339  
 ; LENGTH: 318  
 ; TYPE: PRF  
 ; ORGANISM: Homo sapien  
 US-09-352-616A-339

Query Match 95.1%; Score 1577; DB 3; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5.4e-177;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMPFLLLLPFLYMAAPQIRKMLSSGVCSTVQLPKYVVYTGANTGIGETAKE 60  
 Db 1 MVELMPFLLLLPFLYMAAPQIRKMLSSGVCSTVQLPKYVVYTGANTGIGETAKE 60  
 QY 61 LAORGARVYLACRDVEKELVAKETOTTGNOQVLRKLDLSDTKSIRAWAKGFKAEEKH 120  
 Db 61 LAORGARVYLACRDVEKELVAKETOTTGNOQVLRKLDLSDTKSIRAWAKGFKAEEKH 120  
 QY 121 LHWINNAGVMCMPCYSKTADGFEHIGVNHGFLTHLLLEKLESAPSRIVNSSLAH 180  
 Db 121 LHWINNAGVMCMPCYSKTADGFEHIGVNHGFLTHLLLEKLESAPSRIVNSSLAH 180  
 QY 181 HLGRIHFHNLQGEKFEYNAGLAACHSKLANILFTQELARRIKSGSVTTYSVHPGVOSSELY 240  
 Db 181 HLGRIHFHNLQGEKFEYNAGLAACHSKLANILFTQELARRIKSGSVTTYSVHPGVOSSELY 240  
 QY 241 RHSSFMRMWMLFSPFIKTPOGAQOTRLHCAITGELILSGNHPSDCHVAVMSAQARNET 300  
 Db 241 RHSSFMRMWMLFSPFIKTPOGAQOTRLHCAITGELILSGNHPSDCHVAVMSAQARNET 300  
 QY 301 IARRLMDV 308  
 Db 301 IARRLMDV 308

## RESULT 7

US-09-636-215-339  
 ; Sequence 339, Application US/09636215  
 ; Patent No. 6620922  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Retter, Marc W.  
 ; APPLICANT: Stolk, John A.

APPLICANT: Day, Craig H.  
 APPLICANT: Vedvick, Thomas S.  
 APPLICANT: Carter, Darrick  
 APPLICANT: Li, Samuel  
 APPLICANT: Wang, Aijun  
 APPLICANT: Skeiky, Yasir A.W.  
 APPLICANT: Hepler, William  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
 ; FILE REFERENCE: 210121.4271C17  
 ; CURRENT APPLICATION NUMBER: US/09/636,215  
 ; CURRENT FILING DATE: 2000-08-10  
 ; NUMBER OF SEQ ID NOS: 852  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 339  
 ; LENGTH: 318  
 ; TYPE: PRF  
 ; ORGANISM: Homo sapien  
 US-09-636-215-339

Query Match 95.1%; Score 1577; DB 4; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5.4e-177;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMPFLLLLPFLYMAAPQIRKMLSSGVCSTVQLPKYVVYTGANTGIGETAKE 60  
 Db 1 MVELMPFLLLLPFLYMAAPQIRKMLSSGVCSTVQLPKYVVYTGANTGIGETAKE 60  
 QY 61 LAORGARVYLACRDVEKELVAKETOTTGNOQVLRKLDLSDTKSIRAWAKGFKAEEKH 120  
 Db 61 LAORGARVYLACRDVEKELVAKETOTTGNOQVLRKLDLSDTKSIRAWAKGFKAEEKH 120  
 QY 121 LHWINNAGVMCMPCYSKTADGFEHIGVNHGFLTHLLLEKLESAPSRIVNSSLAH 180  
 Db 121 LHWINNAGVMCMPCYSKTADGFEHIGVNHGFLTHLLLEKLESAPSRIVNSSLAH 180  
 QY 181 HLGRIHFHNLQGEKFEYNAGLAACHSKLANILFTQELARRIKSGSVTTYSVHPGVOSSELY 240  
 Db 181 HLGRIHFHNLQGEKFEYNAGLAACHSKLANILFTQELARRIKSGSVTTYSVHPGVOSSELY 240  
 QY 241 RHSSFMRMWMLFSPFIKTPOGAQOTRLHCAITGELILSGNHPSDCHVAVMSAQARNET 300  
 Db 241 RHSSFMRMWMLFSPFIKTPOGAQOTRLHCAITGELILSGNHPSDCHVAVMSAQARNET 300  
 QY 301 IARRLMDV 308  
 Db 301 IARRLMDV 308

## RESULT 8

US-09-685-166A-339  
 ; Sequence 339, Application US/09685166A  
 ; Patent No. 6630205  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Xu, Jiangchun  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Harlocker, Susan L.  
 ; APPLICANT: Jiang, Yugu  
 ; APPLICANT: Henderson, Robert A.  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Fanger, Gary R.  
 ; APPLICANT: Retter, Marc W.  
 ; APPLICANT: Stolk, John A.  
 ; APPLICANT: Day, Craig H.  
 ; APPLICANT: Vedvick, Thomas S.  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Li, Samuel  
 ; APPLICANT: Wang, Aijun  
 ; APPLICANT: Skeiky, Yasir A.W.  
 ; APPLICANT: Hepler, William  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER



FILE REFERENCE: 210121.427C21  
CURRENT APPLICATION NUMBER: US/09/685,166A  
CURRENT FILING DATE: 2000-10-10  
NUMBER OF SEQ ID NOS: 898  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO: 339  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Homo sapien  
US-09-685-166A-339

Query Match 95.1%; Score 1577; DB 4; Length 318;  
Best Local Similarity 98.7%; Pred. No. 5,4e-177;  
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 MVELMFPILLILPFLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGKETAKE 60
DB 1 MVELMFPILLILPFLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGKETAKE 60
QY 61 LAQGARVYLA CRDVEKELVAKEIOTTGNOQVLRKLDSDTSIRAPAKGFPAEKEK 120
DB 61 LAQGARVYLA CRDVEKELVAKEIOTTGNOQVLRKLDSDTSIRAPAKGFPAEKEK 120
QY 121 LHWINNA GVMWCPYSKTADGEMHIGVNHGFLTHLLEKESAPSRIVVSSLAH 180
DB 121 LHWINNA GVMWCPYSKTADGEMHIGVNHGFLTHLLEKESAPSRIVVSSLAH 180
QY 181 HLGRHFHNLQGEKFNAGLA VCHSKLANILFTQELARLKSGVTTYSVHPGTVOSELV 240
DB 181 HLGRHFHNLQGEKFNAGLA VCHSKLANILFTQELARLKSGVTTYSVHPGTVOSELV 240
QY 241 RHSSFMRMWMLFSFFIKTPQOQAOTRLHCALTGELTSGNHSDDCHVAVWSAQAARNET 300
DB 241 RHSSFMRMWMLFSFFIKTPQOQAOTRLHCALTGELTSGNHSDDCHVAVWSAQAARNET 300
QY 301 IARRLMDV 308
DB 301 IARRLMDV 308
```

#### RESULT 9

US-09-679-426-339

Sequence 339, Application US/09679426

Patent No. 6759515

GENERAL INFORMATION:

APPLICANT: Xu, Jianshun

APPLICANT: Dillon, David C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan L.

APPLICANT: Jiang, Yuqiu

APPLICANT: Henderson, Robert A.

APPLICANT: Kales, Michael D.

APPLICANT: Ranger, Gary R.

APPLICANT: Retter, Marc W.

APPLICANT: Stolk, John A.

APPLICANT: Day, Craig H.

APPLICANT: Vedvick, Thomas S.

APPLICANT: Carter, Darrick

APPLICANT: Li, Samuel

APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Hepler, William

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

FILE REFERENCE: 210121.427C20

CURRENT APPLICATION NUMBER: US/09/679,426

CURRENT FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 895

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO: 339

LENGTH: 318

TYPE: PRT

ORGANISM: Homo sapien

US-09-679-426-339

Query Match

Best Local Similarity 95.1%; Score 1577; DB 4; Length 318;

Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 MVELMFPILLILPFLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGKETAKE 60
DB 1 MVELMFPILLILPFLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGKETAKE 60
QY 61 LAQGARVYLA CRDVEKELVAKEIOTTGNOQVLRKLDSDTSIRAPAKGFPAEKEK 120
DB 61 LAQGARVYLA CRDVEKELVAKEIOTTGNOQVLRKLDSDTSIRAPAKGFPAEKEK 120
QY 121 LHWINNA GVMWCPYSKTADGEMHIGVNHGFLTHLLEKESAPSRIVVSSLAH 180
DB 121 LHWINNA GVMWCPYSKTADGEMHIGVNHGFLTHLLEKESAPSRIVVSSLAH 180
QY 181 HLGRHFHNLQGEKFNAGLA VCHSKLANILFTQELARLKSGVTTYSVHPGTVOSELV 240
DB 181 HLGRHFHNLQGEKFNAGLA VCHSKLANILFTQELARLKSGVTTYSVHPGTVOSELV 240
QY 241 RHSSFMRMWMLFSFFIKTPQOQAOTRLHCALTGELTSGNHSDDCHVAVWSAQAARNET 300
DB 241 RHSSFMRMWMLFSFFIKTPQOQAOTRLHCALTGELTSGNHSDDCHVAVWSAQAARNET 300
QY 301 IARRLMDV 308
DB 301 IARRLMDV 308
```

#### RESULT 10

US-09-270-767-41724

Sequence 41724, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO: 41724

LENGTH: 309

TYPE: PRT

ORGANISM: Drosophila melanogaster

FEATURES:

OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-41724

Query Match

Best Local Similarity 30.9%; Score 512.5; DB 4; Length 309;

Matches 113; Conservative 55; Mismatches 102; Indels 3; Gaps 2;

```
QY 38 QLPGRVVVVTGANTGIGKETAKELAQRGARVYLA CRDVEKELVAKEIOTTGNOQVLR 97
DB 38 QLPGRVVVVTGANTGIGKETAKELAQRGARVYLA CRDVEKELVAKEIOTTGNOQVLR 97
QY 22 XMEGKTVIITGANSGIGKETAKDLAQRGARVYLA CRDVEKELVAKEIOTTGNOQVLR 81
DB 22 XMEGKTVIITGANSGIGKETAKDLAQRGARVYLA CRDVEKELVAKEIOTTGNOQVLR 81
QY 98 KLDLSDTSIRAPAKGFPAEKEKLAHWINNA GVMWCPYSKTA-DGFEHIGVNHGFLTH 156
DB 98 KLDLSDTSIRAPAKGFPAEKEKLAHWINNA GVMWCPYSKTA-DGFEHIGVNHGFLTH 156
QY 82 KLDLSDTSIRAPAKGFPAEKEKLAHWINNA GVMWCPYSKTA-DGFEHIGVNHGFLTH 141
DB 82 KLDLSDTSIRAPAKGFPAEKEKLAHWINNA GVMWCPYSKTA-DGFEHIGVNHGFLTH 141
QY 157 THLLEKESAPSRIVVSSLAHHLGRHFHNLQGEKFNAGLA VCHSKLANILFTQEL 216
DB 157 THLLEKESAPSRIVVSSLAHHLGRHFHNLQGEKFNAGLA VCHSKLANILFTQEL 216
QY 142 THLLIDVLA KRKSAAPARIVVSSLAHHLGRHFHNLQGEKFNAGLA VCHSKLANILFTQEL 201
DB 142 THLLIDVLA KRKSAAPARIVVSSLAHHLGRHFHNLQGEKFNAGLA VCHSKLANILFTQEL 201
QY 217 ARLKSGVTTYSVHPGTVOSELVVRHSSF--KMMWMLFSFFIKTPQOQAOTRLHCALT 274
DB 217 ARLKSGVTTYSVHPGTVOSELVVRHSSF--KMMWMLFSFFIKTPQOQAOTRLHCALT 274
QY 202 AKRLGKTVVNLFLHREMIDSGIWRVPPPLNPPMAITGPFKTKAKAGTIIYATSD 261
DB 202 AKRLGKTVVNLFLHREMIDSGIWRVPPPLNPPMAITGPFKTKAKAGTIIYATSD 261
QY 275 GLEILSGNHSDDCHVAVWSAQAARNETIARRLMD 307
DB 275 GLEILSGNHSDDCHVAVWSAQAARNETIARRLMD 307
```

DB 262 EVANVSGKYMDCKEATINAAAADBEKJIKINE 294

RESULT 11

US-08-545B-2

Sequence 2, Application US/08580545B  
Patent No. 5932713

GENERAL INFORMATION:

APPLICANT: Yoshihisa, Kasukabe

APPLICANT: Koichi, Fujisawa

APPLICANT: Susumu, Nishiguchi

APPLICANT: Yoshiko, Maekawa

APPLICANT: Randy, Allen

TITLE OF INVENTION: COTTON FIBER TISSUE-SPECIFIC GENES

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 601 Thirteenth Street, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/580,545B

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Bretschneider, Barry E.

REGISTRATION NUMBER: 28,055

REFERENCE/DOCKET NUMBER: 04473/068001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202/783-5070

TELEFAX: 202/783-2331

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 323 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-545B-2

Query Match 29.8%; Score 494.5; DB 2; Length 323;  
Best Local Similarity 40.1%; Pred. No. 2.5e-49;  
Matches 110; Conservative 53; Mismatches 100; Indels 11; Gaps 4;

DB 45 VTGANTGIGKETAKELAQGARYVYACRDVEKGEIVAKXIQTGGNQVLVRKLDSDT 104

DB 38 LVTGASGIGLETSRVYALRGVYIIGARNMKANAKKIVENPRARIDVLELDIST 97

DB 105 KSIRAMAKGFKAEEKLHWMINNAGVMPCYKTDGFEHIGVNHIGFLLTHLLEKL 164

DB 98 NSIRSFADNFIALHLPINIIINAGIMFCFQLSQNGLEVOFATNHIGFLLTNLLDITM 157

DB 165 KESAPS-----RIVNVSSLAH---HGRHFHNLQGEKFTYNAGLAYCHSKLANILFTQEL 216

DB 158 KNTVKATGIGRNVVNISSIAHNYCYKKGIRFEKINDKQGYSEKRAVIGOSKLANILHANEL 217

DB 217 ARLKSGV--TIVSVHPTVQSELVHSSFMKMMWLFSFPI-KTPOGAQTRLHCALT 273

DB 218 SRLQEGVNIITVNSVHPTLITPLFRHSADLMKLLKFFSFLMKVDPQGAATTCVVALH 277

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

Sequence 2, Application US/09262653A  
Patent No. 616294

GENERAL INFORMATION:

APPLICANT: Yoshihisa, Kasukabe

APPLICANT: Koichi, Fujisawa

APPLICANT: Susumu, Nishiguchi

APPLICANT: Yoshiko, Maekawa

APPLICANT: Randy, Allen

TITLE OF INVENTION: COTTON FIBER TISSUE-SPECIFIC GENES

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 601 Thirteenth Street, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/262,653A

FILING DATE:

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Bretschneider, Barry E.

REGISTRATION NUMBER: 28,055

REFERENCE/DOCKET NUMBER: 04473/068001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202/783-5070

TELEFAX: 202/783-2331

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 323 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-262-653A-2

Query Match 29.8%; Score 494.5; DB 3; Length 323;  
Best Local Similarity 40.1%; Pred. No. 2.5e-49;  
Matches 110; Conservative 53; Mismatches 100; Indels 11; Gaps 4;

DB 45 VTGANTGIGKETAKELAQGARYVYACRDVEKGEIVAKXIQTGGNQVLVRKLDSDT 104

DB 38 LVTGASGIGLETSRVYALRGVYIIGARNMKANAKKIVENPRARIDVLELDIST 97

DB 105 KSIRAMAKGFKAEEKLHWMINNAGVMPCYKTDGFEHIGVNHIGFLLTHLLEKL 164

DB 98 NSIRSFADNFIALHLPINIIINAGIMFCFQLSQNGLEVOFATNHIGFLLTNLLDITM 157

DB 165 KESAPS-----RIVNVSSLAH---HGRHFHNLQGEKFTYNAGLAYCHSKLANILFTQEL 216

DB 158 KNTVKATGIGRNVVNISSIAHNYCYKKGIRFEKINDKQGYSEKRAVIGOSKLANILHANEL 217

DB 217 ARLKSGV--TIVSVHPTVQSELVHSSFMKMMWLFSFPI-KTPOGAQTRLHCALT 273

DB 218 SRLQEGVNIITVNSVHPTLITPLFRHSADLMKLLKFFSFLMKVDPQGAATTCVVALH 277

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

DB 278 PRLNGVTGKYFADCNEMRPSVARNBSLGRLEWE 311

DB 274 EGEILISGNHFSCHVAWVSAQARNETIARLMD 307

```

FILE REFERENCE: P2031P1
CURRENT APPLICATION NUMBER: US/09/489,847
CURRENT FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
EARLIER FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO: 182
LENGTH: 330
TYPE: PRT
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: SITE
LOCATION: (247)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-489-847-182

```

Query Match 28.6%; Score 474; DB 4; Length 330;

Best Local Similarity 39.6%; Pred. No. 6.7e-47; Matches 114; Conservative 43; Mismatches 97; Indels 34; Gaps 7;

```

QY 40 PGKVVVVGANTGIGKETAKELAORGARVYLACDVKELVAKEIQTGGNOQVLRKL 99
DB 42 PDRVAIVTGTGDTGIGYSAKHLARLGNHVIAGNNDKAKQVYSKIEEPLNDKVELYC 101
QY 100 DLSQTSIRAWAKGFKAEKHLHWINNAGVWCPSYKTDAGFEMHIGVNLGHFLITHL 159
DB 102 DLAHTSTIRQVQCFKMKKIPDLHVLINNAGVWMPQKTRDGFEEHGLVYLGFFLITNL 161
QY 160 LLEKLKESAE-----PSRIYVSSLAHGRIFHNLIQGEKFPYNAGLACHSKLNIPTQES 215
DB 162 LLDLTKEGSGPCHSRVTVSSATHYVAELNMDLQSSACYSFAAAYAQSKLALVLTTH 221
QY 216 LARRL--KSGVTVTVSVHPTVQSELYVHSSFRMWM-----WWLFSFFIKTPQOGA 264
DB 222 LQRLLAAGSHVTANVDPGVVNTDXYKH---VFMAIRLAKKLIGWLLF---KTPDEGA 274
QY 265 QTRHLCAITGELTILSGNHF-----SDCHVAMVSAQARNETIARRLM 306
DB 275 WTSIYAATVPELGEVGRVLYNEKETKSLHVTY-----NQKLOOQJLM 316

```

RESULT 14

US-08-530-165-7

```

Sequence 7, Application US/08530165
Patent No. 5807081
GENERAL INFORMATION:
APPLICANT: Isaac, Peter G.
APPLICANT: Roberts, Jeremy A.
APPLICANT: Coupe, Simon A.
TITLE OF INVENTION: Control of Plant Abscission and Pod Dehiscence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: 20005
ZIP: USA

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,165
FILING DATE: 29-SEP-1995
CLASSIFICATION: 800
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/00689
FILING DATE: 31-MAR-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9306726.2
FILING DATE: 31-MAR-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2500
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-530-165-7

```

Query Match 28.1%; Score 466.5; DB 2; Length 320;

Best Local Similarity 38.9%; Pred. No. 4.9e-46; Matches 111; Conservative 48; Mismatches 115; Indels 11; Gaps 4;

```

QY 34 TSTVQLPGKVVVVGANTGIGKETAKELAORGARVYLACDVKELVAKEIQTGGNOQ 93
DB 26 TSHDAKHITAITIGTSGIGLEAARVLGMRGAHYIISRNKKAANDSKEMILQVYPNAR 85
QY 94 VLVRLKLDSTYSIRAWAKGFKAEKHLHWINNAGVWCPSYKTDAGFEMHIGVNLGH 153
DB 86 ICDGLDSSIKSVSFIHQFLANVPLNIIINNAGVWCPQLSEDEGIESQFATNHIGH 145
QY 154 FLLTHLLEKLKESAE-----PSRIYVSSLAH---HGRIFHNLIQGEKFPYNAGLACHS 205
DB 146 FLLTHLLEKLKESAE-----PSRIYVSSLAH---HGRIFHNLIQGEKFPYNAGLACHS 205
QY 206 KLANILFTOLARPKSGCV--TVYSVHPTVQSELYVHSSFRMWMWLFSEFI-KTPQO 262
DB 206 KLANILHSNALRSKIQEBSVNTITNSVHGLITTLFPHSGGLVAMVLAAMSGFLKKNIPQ 265
QY 263 GAQTRHLCAITGELTILSGNHFSDCVAVWSAQAARNETIARRLM 307
DB 266 GAATTCYVALHPDLKDVTKGYFADCNVTPPSNFATDTTLADKLWD 310

```

RESULT 15

US-09-489-847-330

```

Sequence 330, Application US/09489847
Patent No. 6476195
GENERAL INFORMATION:
APPLICANT: Rosen et al
APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031P1
CURRENT APPLICATION NUMBER: US/09/489,847
CURRENT FILING DATE: 2000-01-24
EARLIER APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05
EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
EARLIER FILING DATE: 1998-08-06

```

```

: NUMBER OF SEQ ID NOS: 376
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 330
: LENGTH: 330
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (38)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (247)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-489-847-330

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 18:01:08 ; Search time 156 Seconds  
(without alignments)  
726.657 Million cell updates/sec

Title: US-10-021-002-2  
Perfect score: 1659  
Sequence: 1 MEIEMFPLLLPLFLYMA.....RNETIARLMDVIVTCMASQ 316

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A\_Geneseq\_23Sep04.\*  
2: geneseqp1980s.\*  
3: geneseqp1990s.\*  
4: geneseqp2000s.\*  
5: geneseqp2001s.\*  
6: geneseqp2002s.\*  
7: geneseqp2003as.\*  
8: geneseqp2003bs.\*  
9: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1652	99.6	316	2 AAW03198	AAW03198 Prostate
2	1577	95.1	318	2 AAY82014	AAY82014 Human imm
3	1577	95.1	318	3 AAB12158	AAB12158 Hydrophob
4	1577	95.1	318	4 AAM01129	AAM01129 Human pro
5	1577	95.1	318	4 AAM93777	AAM93777 Human pol
6	1577	95.1	318	4 AAU69774	AAU69774 Human pro
7	1577	95.1	318	4 AAB74812	AAB74812 Prostate
8	1577	95.1	318	4 AAG99014	AAG99014 Human pro
9	1577	95.1	318	4 ABU71665	ABU71665 Prostate
10	1577	95.1	318	5 ABG96558	ABG96558 Human sho
11	1577	95.1	318	5 ABR95334	ABR95334 Human P50
12	1577	95.1	318	6 ABR54346	ABR54346 Prostate
13	1577	95.1	318	7 ADB75403	ADB75403 Prostate
14	1577	95.1	318	7 ADB13789	ADB13789 Human pro
15	1577	95.1	318	7 ADG26205	ADG26205 Human pro
16	1577	95.1	318	7 ADJ71200	ADJ71200 Human hea
17	1577	95.1	318	8 ADI06507	ADI06507 Human tum
18	1577	95.1	318	8 ADI31755	ADI31755 Human pro
19	1577	95.1	333	8 AAY48616	AAY48616 Human bre
20	1573	94.8	318	7 ABR63833	ABR63833 Human oxi
21	1573	94.8	318	7 ABR82986	ABR82986 Human CGI
22	1555	93.7	314	8 AAB36900	AAB36900 Human ARS
23	1169	70.1	248	8 ADH45298	ADH45298 Human enz
24	1130	68.1	316	6 ADA54192	ADA54192 Human pro
25	1130	68.1	316	7 ABR82985	ABR82985 Human umm

26	1129	68.1	316	5 AAE14438	AAE14438 Human dru
27	1117.5	67.4	321	5 ABP69268	ABP69268 Human pol
28	1112	67.0	260	7 ADG76361	ADG76361 Human inc
29	1020	61.5	278	6 ABUS2578	ABUS2578 Human NOV
30	806	48.6	246	4 AAB73685	AAB73685 Human oxi
31	728	43.9	300	4 ABR58826	ABR58826 Drosophi
32	726	43.8	296	4 ABR58831	ABR58831 Drosophi
33	706.5	42.6	331	4 AAB88356	AAB88356 Human mem
34	705.5	42.5	331	3 AAY93722	AAY93722 Human PRO
35	705.5	42.5	331	4 AAB66121	AAB66121 Protein o
36	705.5	42.5	331	4 AAE05174	AAE05174 Human dru
37	705.5	42.5	331	6 AAB37278	AAB37278 Human gen
38	705.5	42.5	331	6 ABO33614	ABO33614 Novel hum
39	705.5	42.5	331	7 ABO44467	ABO44467 Human sec
40	705.5	42.5	331	7 ABO33491	ABO33491 Novel hum
41	705.5	42.5	331	7 ADC17985	ADC17985 Human PRO
42	705.5	42.5	331	7 ADD70631	ADD70631 Human sec
43	705.5	42.5	331	7 ADD39708	ADD39708 Human sec
44	705.5	42.5	331	7 ADD70154	ADD70154 Human sec
45	705.5	42.5	331	7 ADD38275	ADD38275 Human sec

## ALIGNMENTS

RESULT 1  
ID AAW03198  
AAW03198 standard; protein; 316 AA.

AC AAW03198;  
XX

DT 26-FEB-1997 (first entry)  
XX

DE Prostate specific reductase.  
XX

KM Prostate specific reductase; human; prostate; prostate cancer; inhibitor;  
XX

KW antagonist; cancer metastasis; prostate disorder; antibody; therapy; PSR;  
XX

OS prostate cancer vaccine.  
XX

OS Homo sapiens.  
XX

PN WO9622360-A1.  
XX

PD 25-JUL-1996.  
XX

PF 20-JAN-1995; 95WO-US001827.  
XX

PR 20-JAN-1995; 95WO-US001827.  
XX

PA (HUMA-) HUMAN GENOME SCI INC.  
XX

PI He W, Meisner PS, Hudson PL, Rosen CA;  
XX

DR WPI, 1996-354517/35.  
XX

DR N-PSDB; AAT13121.  
XX

PT Human prostate specific reductase - useful for diagnosing and treating  
XX

PS prostate cancer and screening new antagonists.  
XX

PS Claim 19; Page 44-45; 57pp; English.  
XX

This sequence represents the human prostate specific reductase (PSR). PSR genes are present in all cells of the body, however transcription and expression of PSR is limited to the prostate in normal individuals. When prostate cancer is present, the cancer cells migrate and cause other cells to begin expressing PSR. It is therefore thought that the presence of mRNA encoding PSR in cells of the host (other than those derived from the prostate) is indicative of prostate cancer metastases. The expression of PSR provides a diagnostic marker for prostate disorders. The DNA encoding this sequence can therefore be used to diagnose metastasis of prostate cancer cells. PSR may be used to screen for PSR antagonists, which can then be used as inhibitors of PSR expression. Antibodies against this protein sequence can be used to target prostate cells, and

CC could also be used as part of a prostate cancer vaccine  
 XX Sequence 316 AA;

Query Match 99.6%; Score 1652; DB 2; Length 316;  
 Best Local Similarity 99.7%; Pred. No. 9.1e-163;  
 Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVELMFPLLLLLPFLYMAAPQIRKMLSSGVCSTVQLPGKVVVVGANTGIGKETAKE 60  
 DB 1 MVELMFPLLLLLPFLYMAAPQIRKMLSSGVCSTVQLPGKVVVVGANTGIGKETAKE 60  
 QY 61 LAQSGARYYLAACRDVEKGEIVAKEIOTTGNOQLVYRKLDLSDTKSIRAWAKGFKAEEKH 120  
 DB 61 LAQSGARYYLAACRDVEKGEIVAKEIOTTGNOQLVYRKLDLSDTKSIRAWAKGFKAEEKH 120  
 QY 121 LHWNNAGVVMGVCYSTADGFEMHIGVNHGHLTHLLLEKLEKESAPRIVVSSLAH 180  
 DB 121 LHWNNAGVVMGVCYSTADGFEMHIGVNHGHLTHLLLEKLEKESAPRIVVSSLAH 180  
 QY 181 HLGRHFNHLOGEKFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 DB 181 HLGRHFNHLOGEKFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 QY 241 RHSSFMKMMWTLFSPFKTPOQGAQTRLHCALTEGLEILSGNHPSDCHVAVWSAQARNET 300  
 DB 241 RHSSFMKMMWTLFSPFKTPOQGAQTRLHCALTEGLEILSGNHPSDCHVAVWSAQARNET 300  
 QY 301 IARLMDVITVCMASQ 316  
 DB 301 IARLMDVITVCMASQ 316

RESULT 2  
 AAY82014  
 ID AAY82014 standard; protein; 318 AA.

AC AAY82014;  
 DT 13-JUN-2000 (first entry)  
 XX Human immunogenic prostate tumour protein sequence SEQ ID NO:339.  
 DE Human immunogenic prostate tumour protein sequence SEQ ID NO:339.  
 XX Human; prostate cancer; diagnosis; tumour; gene therapy; detection;  
 KW immunogenic; cytostatic; vaccine.  
 XX Homo sapiens.

PN WO200004149-A2.

XX 27-JAN-2000.

PF 14-JUL-1999; 99WO-US015638.

XX 14-JUL-1998; 98US-00115453.

PR 14-JUL-1998; 98US-00116134.

PR 23-SEP-1998; 98US-00158812.

PR 15-JAN-1999; 99US-00232149.

PR 09-APR-1999; 99US-00288946.

XX (CORI-) CORIXA CORP.

PI DILLON DC, Harlocker SL, Yudin J, Xu J, Mitcham JL;

DR WPI, 2000-171268/15.

XX New polypeptide useful for treating and diagnosing prostate cancer  
 PT comprises an immunogenic portion of prostate tumor protein.  
 XX Claim 3; Page 209-210; 263pp; English.

CC The present invention describes isolated polypeptides, comprising an  
 CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides  
 CC and polynucleotides encoding them have cytostatic activity and can be  
 CC used in vaccines and in gene therapy. The polypeptides and  
 CC polynucleotides encoding them, antigen presenting cells which express the  
 CC polypeptides, antibodies against the polypeptides and vaccines comprising  
 CC them can be used for inhibiting the development of prostate cancer in a  
 CC patient. The polypeptides can be used to generate antibodies or anti-  
 CC idiotypic antibodies for passive immuno therapy. A portion of the  
 CC polynucleotides encoding the polypeptides can be used as a probe or to  
 CC modulate the expression of the polypeptides. AA062241 to AA066591 and  
 CC AAY82000 to AAY82020 represent sequences used in the exemplification of  
 CC the present invention

Sequence 318 AA;

Query Match 95.1%; Score 1577; DB 3; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5.9e-155;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMFPLLLLLPFLYMAAPQIRKMLSSGVCSTVQLPGKVVVVGANTGIGKETAKE 60  
 DB 1 MVELMFPLLLLLPFLYMAAPQIRKMLSSGVCSTVQLPGKVVVVGANTGIGKETAKE 60  
 QY 61 LAQSGARYYLAACRDVEKGEIVAKEIOTTGNOQLVYRKLDLSDTKSIRAWAKGFKAEEKH 120  
 DB 61 LAQSGARYYLAACRDVEKGEIVAKEIOTTGNOQLVYRKLDLSDTKSIRAWAKGFKAEEKH 120  
 QY 121 LHWNNAGVVMGVCYSTADGFEMHIGVNHGHLTHLLLEKLEKESAPRIVVSSLAH 180  
 DB 121 LHWNNAGVVMGVCYSTADGFEMHIGVNHGHLTHLLLEKLEKESAPRIVVSSLAH 180  
 QY 181 HLGRHFNHLOGEKFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 DB 181 HLGRHFNHLOGEKFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 QY 241 RHSSFMKMMWTLFSPFKTPOQGAQTRLHCALTEGLEILSGNHPSDCHVAVWSAQARNET 300  
 DB 241 RHSSFMKMMWTLFSPFKTPOQGAQTRLHCALTEGLEILSGNHPSDCHVAVWSAQARNET 300  
 QY 301 IARLMDV 308  
 DB 301 IARLMDV 308

RESULT 3  
 AAB12158  
 ID AAB12158 standard; protein; 318 AA.

XX AAB12158;

DT 02-FEB-2001 (first entry)

XX Hydrophobic domain protein isolated from HT-1080 cells.

DE Human; secreted protein; membrane protein; hydrophobic domain;

KW proliferation control; differentiation induction; material transport;

KW biophysics; signal receptor; ion channel; transporter; immunostimulant;

KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;

KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;

KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer.

XX Homo sapiens.

XX WO200029448-A2.

XX 25-MAY-2000.

XX 17-NOV-1999; 99WO-JP006412.

XX 17-NOV-1998; 98JP-00326255.

XX 22-DEC-1998; 98JP-00364315.

XX 16-MAR-1999; 99JP-00069811.

27-APR-1999; 99JP-00119299.  
 19-MAY-1999; 99JP-00138169.  
 (SAGA ) SAGAMI CHEM RES CENT.  
 (PROT-) PROTEGENE INC.  
 Kato S, Kimura T;  
 WPI; 2000-387753/33.  
 N-PSDB; AAA62061, AAA62071.  
 proteins comprising hydrophobic regions, such as secretory and membrane proteins, useful in research and diagnostics and having various activities e.g. immunomodulatory, antiinflammatory, chemokinetic, hemostatic, thrombolytic.  
 Claim 1; Page 369-370; 410pp; English.  
 Secretory proteins play important roles in the proliferation control, the differentiation induction, the material transport and the biophysics of cells. Membrane proteins have important roles as signal receptors, ion channels and transporters. The present sequence is a human protein which has at least one hydrophobic domain. This protein may be a secretory or a membrane protein. The present protein may have cytokine and cell proliferation/differentiation activity, immune stimulating or suppressing activity, haemopoiesis activity, tissue growth activity, haemostatic activity/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, anti-inflammatory activity and tumor inhibition activity. The present protein could therefore be used for treatment of autoimmune disease, Alzheimer's disease, Parkinson's disease, and cancer  
 Sequence 318 AA:  
 Query Match 95.1%; Score 1577; DB 3; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5.9e-155;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MVELMPELLLLLPFLVMAAPQIRKMLSSGVTSTVQPGKVVVVTGANTGIGKETAKE 60  
 DB 1 MVELMPELLLLLPFLVMAAPQIRKMLSSGVTSTVQPGKVVVVTGANTGIGKETAKE 60  
 QY 61 LAORGARVYLA CRDVEKELVAKEIQTTGNQOVLVRKLDLSDTKSIRAMAKGFABEEKH 120  
 DB 61 LAORGARVYLA CRDVEKELVAKEIQTTGNQOVLVRKLDLSDTKSIRAMAKGFABEEKH 120  
 QY 121 LHWANNAGVMMCPYSKTADGFEHMGVNLGHFLTHLLLEKESAPSRIVNSSLAH 180  
 DB 121 LHWANNAGVMMCPYSKTADGFEHMGVNLGHFLTHLLLEKESAPSRIVNSSLAH 180  
 QY 181 HLGRIFHNLQGEKFNAGLAYCHSKLANILFTQELARLKSGVTTYSVHPGTQSELY 240  
 DB 181 HLGRIFHNLQGEKFNAGLAYCHSKLANILFTQELARLKSGVTTYSVHPGTQSELY 240  
 QY 241 RHSSFMKMMWMLFSPFIKTPOOGAOTRLHCALTBSLETLISGNHPSDCVAVWSAQRNET 300  
 DB 241 RHSSFMKMMWMLFSPFIKTPOOGAOTRLHCALTBSLETLISGNHPSDCVAVWSAQRNET 300  
 QY 301 IARRLMVDV 308  
 DB 301 IARRLMVDV 308  
 RESULT 4  
 AA001129  
 ID AA001129 standard; protein; 318 AA.  
 AC AA001129;  
 XX  
 DT 04-OCT-2001 (first entry)  
 XX  
 DE Human prostate-specific amino acid sequence P509S.  
 XX

Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
 cytostatic; gene therapy; metastasis.  
 Homo sapiens.  
 WO200151633-A2.  
 19-JUL-2001.  
 16-JAN-2001; 2001WO-US001574.  
 14-JAN-2000; 2000US-00483672.  
 (CORI-) CORIXA CORP.  
 Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
 Kalos MD, Fanger GR, Day CH, Rector MM, Stolk JA, Skeiky YAM;  
 Wang A, Yeager MJ;  
 WPI; 2001-425873/45.  
 New polynucleotide encoding a prostate-specific protein, for diagnosing, monitoring and treating prostate cancer in a patient and for use in vaccines.  
 Claim 2; Page 335; 543pp; English.  
 The present invention describes polynucleotide sequences (I) which encode prostate-specific proteins (II). (I) and (II) have cytostatic activity, and can be used in vaccine production and gene therapy. (I), (II), antibodies to (I) or (II) are used to treat cancer in a patient. (I) and (II) are also used in the detection of cancer in a patient. The cancer that is diagnosed or treated is particularly prostate cancer. (I) and (II) can be used in vaccines. The antibodies or (I) and (II) can also be used to improve diagnostic and therapeutic methods for prostate cancer. They can indicate the level of metastasis as well as the prostate volume. AA03357 to AA03944 and AA00115 to AA00118 represent polynucleotide and amino acid sequences used in the exemplification of the present invention  
 Sequence 318 AA:  
 Query Match 95.1%; Score 1577; DB 4; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5.9e-155;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MVELMPELLLLLPFLVMAAPQIRKMLSSGVTSTVQPGKVVVVTGANTGIGKETAKE 60  
 DB 1 MVELMPELLLLLPFLVMAAPQIRKMLSSGVTSTVQPGKVVVVTGANTGIGKETAKE 60  
 QY 61 LAORGARVYLA CRDVEKELVAKEIQTTGNQOVLVRKLDLSDTKSIRAMAKGFABEEKH 120  
 DB 61 LAORGARVYLA CRDVEKELVAKEIQTTGNQOVLVRKLDLSDTKSIRAMAKGFABEEKH 120  
 QY 121 LHWANNAGVMMCPYSKTADGFEHMGVNLGHFLTHLLLEKESAPSRIVNSSLAH 180  
 DB 121 LHWANNAGVMMCPYSKTADGFEHMGVNLGHFLTHLLLEKESAPSRIVNSSLAH 180  
 QY 181 HLGRIFHNLQGEKFNAGLAYCHSKLANILFTQELARLKSGVTTYSVHPGTQSELY 240  
 DB 181 HLGRIFHNLQGEKFNAGLAYCHSKLANILFTQELARLKSGVTTYSVHPGTQSELY 240  
 QY 241 RHSSFMKMMWMLFSPFIKTPOOGAOTRLHCALTBSLETLISGNHPSDCVAVWSAQRNET 300  
 DB 241 RHSSFMKMMWMLFSPFIKTPOOGAOTRLHCALTBSLETLISGNHPSDCVAVWSAQRNET 300  
 QY 301 IARRLMVDV 308  
 DB 301 IARRLMVDV 308





Db 1 MVELMPFLLILLPFLLYMAAPQIRKMLSSGVCSTVQLPKGVVVTGANTGICKETAKE 60  
QY 61 LAORGARVYLACRDEKELVAKETIOTTGNQVLRKLDLSDTKSIRAMAKGFABEKH 120  
Db 61 LAORGARVYLACRDEKELVAKETIOTTGNQVLRKLDLSDTKSIRAFAGFLABEKH 120  
QY 121 LHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLLEKESAPSRIVVSSLAH 180  
Db 121 LHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLLEKESAPSRIVVSSLAH 180  
QY 181 HLGRIHFHNLOGEKRYNAGLAACHSKLANILFTQELARLKSGSVTTYSVHPGTVOSELY 240  
Db 181 HLGRIHFHNLOGEKRYNAGLAACHSKLANILFTQELARLKSGSVTTYSVHPGTVOSELY 240  
QY 241 RHSSFRMMWMLFSFFIKTPQOGAQTSLHCAITGELTSLGSHFSDCHAVAVSAQARNET 300  
Db 241 RHSSFRMMWMLFSFFIKTPQOGAQTSLHCAITGELTSLGSHFSDCHAVAVSAQARNET 300  
QY 301 IARRLMDV 308  
Db 301 IARRLMDV 308

RESULT 7  
AAB74812  
ID AAB74812 standard; protein, 318 AA.  
AC AAB74812;  
AD AAB74812;  
AE 14-JUN-2001 (first entry)  
AF Prostate tumour antigen predicated amino acid sequence for P509S.  
AG Prostate tumour antigen; prostate tumour; therapy; diagnosis;  
AH Human; prostate cancer; immunogenic; cytosstatic; vaccine.  
AI Homo sapiens.  
AJ WO200125272-A2.  
AK 12-APR-2001.  
AL 04-OCT-2000; 2000WO-US027464.  
AM 04-OCT-1999; 99US-0157455P.  
AN (CORI-) CORIXA CORP.  
AO Xu J, Skeiky YAM, Reed SG, Cheever MA;  
AP WPI; 2001-245062/25.  
AQ N-PSDB; AAH02742.  
AR Prostate specific protein and its encoding polynucleotide, useful for the  
AS treatment and diagnosis of prostate cancer.  
AT Claim 3; Page 220-221; 276pp; English.  
AX The present invention describes an isolated polypeptide (I) comprising at  
AY least an immunogenic portion of a prostate tumour antigen protein or its  
AZ variant. (I) have cytostatic activity and can be used in vaccine  
BA production. (I), prostate tumour antigen polynucleotides, an antigen  
BB presenting cell (APC e.g. a dendritic cell) that expresses (I), and a  
BC pharmaceutical composition containing (I) are useful for inhibiting the  
BD development of cancer in a patient. Antibodies specific for prostate  
BE specific proteins and oligonucleotides that hybridise to a polynucleotide  
BF that encodes a prostate specific protein are useful for detecting the  
BG presence or absence of a cancer or monitoring the progression the  
BH progression of a cancer, especially prostate cancer. AAH02422 to AAH2872,  
BI AAB74798 to AAB74821 and AAB74830 are sequences used in the  
BJ exemplification of the present invention

SQ Sequence 318 AA;  
Query Match 95.1%; Score 1577; DB 4; Length 318;  
Best Local Similarity 98.7%; Pred. No. 5.9e-155;  
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MVELMPFLLILLPFLLYMAAPQIRKMLSSGVCSTVQLPKGVVVTGANTGICKETAKE 60  
Db 1 MVELMPFLLILLPFLLYMAAPQIRKMLSSGVCSTVQLPKGVVVTGANTGICKETAKE 60  
QY 61 LAORGARVYLACRDEKELVAKETIOTTGNQVLRKLDLSDTKSIRAMAKGFABEKH 120  
Db 61 LAORGARVYLACRDEKELVAKETIOTTGNQVLRKLDLSDTKSIRAFAGFLABEKH 120  
QY 121 LHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLLEKESAPSRIVVSSLAH 180  
Db 121 LHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLLEKESAPSRIVVSSLAH 180  
QY 181 HLGRIHFHNLOGEKRYNAGLAACHSKLANILFTQELARLKSGSVTTYSVHPGTVOSELY 240  
Db 181 HLGRIHFHNLOGEKRYNAGLAACHSKLANILFTQELARLKSGSVTTYSVHPGTVOSELY 240  
QY 241 RHSSFRMMWMLFSFFIKTPQOGAQTSLHCAITGELTSLGSHFSDCHAVAVSAQARNET 300  
Db 241 RHSSFRMMWMLFSFFIKTPQOGAQTSLHCAITGELTSLGSHFSDCHAVAVSAQARNET 300  
QY 301 IARRLMDV 308  
Db 301 IARRLMDV 308

RESULT 8  
AAG99014  
ID AAG99014 standard; protein, 318 AA.  
AC AAG99014;  
AD AAG99014;  
AE 25-SEP-2001 (first entry)  
AF Human prostate-specific amino acid sequence P509S.  
AG Human prostate-specific amino acid sequence P509S.  
AH Human; prostate cancer; therapy; diagnosis; cat eye syndrome;  
AI chromosome 22q11.2; prostate-specific protein; chromosome 1;  
AJ prostate specific antigen; PSA.  
AK Homo sapiens.  
AL WO200134802-A2.  
AM 17-MAY-2001.  
AN 09-NOV-2000; 2000WO-US030904.  
AO 12-NOV-1999; 99US-00439313.  
AP 18-NOV-1999; 99US-00443686.  
AQ (CORI-) CORIXA CORP.  
AR Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
AS Kaios MD, Retter MW, Scolk JA, Day CH, Skeiky YAM, Wang A;  
AT WPI; 2001-308785/32.  
AX Isolated polypeptide comprising at least an immunogenic portion of a  
AY prostate-specific protein, useful in the diagnosis and therapy of  
AZ prostate cancer.  
BA Claim 3; Page 234-235; 325pp; English.  
BB The present invention describes an isolated polypeptide (P1) comprising  
BC at least an immunogenic portion of a prostate-specific protein, or its  
BD variant. Also described are polynucleotides (N1) encoding (P1) and  
BE (N1) have cytostatic activity and can be used in vaccine production. The

CC polypeptides, nucleic acids and antibodies from the present invention are  
 CC useful in the diagnosis and therapy of prostate cancer. Prostate specific  
 CC genes P704P, P712P, P774P, P775P and B305D are located in a genomic  
 CC region on chromosome 22q11.2 known as the Cat Eye Syndrome region.  
 CC Prostate specific antigen (PSA) P501S was located on chromosome 1.  
 CC AA84671 to AA85143 and AAG9800 to AAG9977 represent polynucleotide  
 CC and polypeptide sequences used in the exemplification of the present  
 CC invention

CC Sequence 318 AA;

Query Match 95.1%; Score 1577; DB 4; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5.9e-155;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMFPLLILLPFLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGKETAKE 60  
 DB 1 MVELMFPLLILLPFLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGKETAKE 60  
 QY 61 LAQGARVYLACRVEKGEIVAKEIOTTGNQOVIVRKLDLSDTKSIRAMAKGFAEEKH 120  
 DB 61 LAQGARVYLACRVEKGEIVAKEIOTTGNQOVIVRKLDLSDTKSIRAMAKGFAEEKH 120  
 QY 121 LHVINNAGVMMCPYSKTADGPEMHIGVNLGHFLTHLLLEKESASPRIVNSSLAH 180  
 DB 121 LHVINNAGVMMCPYSKTADGPEMHIGVNLGHFLTHLLLEKESASPRIVNSSLAH 180  
 QY 181 HIGRIHFHNLQGEKFNAGLAICHSKLANILFTQELARLKSGGVTTSVHGTVQSELY 240  
 DB 181 HIGRIHFHNLQGEKFNAGLAICHSKLANILFTQELARLKSGGVTTSVHGTVQSELY 240  
 QY 241 RHSSFRMWMWLFSPFIKTPOQAGQVRLHCAITGELISGNHFSQCHVAVTSAQARNET 300  
 DB 241 RHSSFRMWMWLFSPFIKTPOQAGQVRLHCAITGELISGNHFSQCHVAVTSAQARNET 300  
 QY 301 IARLMDV 308  
 DB 301 IARLMDV 308

RESULT 9

ID ABU71665 standard; protein; 318 AA.

XX ABU71665;  
 XX 10-JUN-2003 (first entry)  
 DE Prostate cancer associated protein #6.  
 XX Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
 KM immunogen; cancer; prostate specific antigen; PSA;  
 KM prostatic acid phosphatase; PAP; prostate specific membrane antigen;  
 XX PMA.  
 OS Homo sapiens.  
 XX US2002192763-A1.  
 PD 19-DEC-2002.  
 XX 29-JUN-2001; 2001US-00895793.  
 PF 04-OCT-1999; 99US-0157455P.  
 PR 04-OCT-2000; 2000US-00679272.  
 PR 28-MAR-2001; 2001US-00822827.  
 XX (XUJ/) XU J.  
 PA (DILL/) DILLON D. C.  
 PA (MITC/) MITCHAM J. L.  
 PA (HARL/) HARLOCKER S. L.  
 PA (JIANG/) JIANG Y.  
 PA (KALOS/) KALOS M. D.

PA (FANG/) FANGER G. R.  
 PA (RETT/) RETTER M. W.  
 PA (STOL/) STOLK J. A.  
 PA (DAYC/) DAY C. H.  
 PA (VEDV/) VEDVICK T. S.  
 PA (CART/) CARTER D.  
 PA (LISX/) LI S. X.  
 PA (WANG/) WANG A.  
 PA (SKEI/) SKEIKY Y. A. W.  
 PA (HEPL/) HEPLER M. T.  
 PA (HEND/) HENDERSON R. A.  
 PA (HURA/) HURAL J.  
 PA (MCNE/) MCNEILL P. D.  
 PA (HOUG/) HOUGHTON R. L.  
 PA (DEAS/) Y. DE BASSOLS C. V.  
 PA (FOYT/) FOY T. M.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 Li SX, Wang A, Skeiky YAM, Hepler MT, Henderson RA, Hural J;  
 Mcneill PD, Houghton RL, Y De Bassolciv, Foy TM;  
 WPI, 2001-245062/25.

Prostate specific protein and its encoding polynucleotide, useful for the  
 treatment and diagnosis of prostate cancer.

Example 1; SEQ ID NO 339; 85pp; English.

The invention describes a fusion protein comprising at least one amino  
 acid sequence of immunogenic portions of any of the 3 sequences not  
 defined in the specification, or sequences having at least 70 or 90 %  
 sequence identity to any one of the 3 sequences defined in the USPTO web  
 site, which is encoded by any of the 4 nucleotide sequences not defined  
 in the specification. The fusion protein, composition and methods are  
 useful for diagnosing, preventing and/or treating cancer, particularly  
 prostate cancer. The proteins are useful as markers to indicate the  
 presence or absence of cancer. This is the amino acid sequence of a  
 prostate cancer therapy associated protein. Note: The sequence data for  
 this patent did not form part of the printed specification, but was  
 obtained in electronic format directly from the US patent office at  
 seqdata.uspto.gov/sequence.html?docid=US20020192763

Sequence 318 AA;

Query Match 95.1%; Score 1577; DB 4; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5.9e-155;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMFPLLILLPFLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGKETAKE 60  
 DB 1 MVELMFPLLILLPFLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGKETAKE 60  
 QY 61 LAQGARVYLACRVEKGEIVAKEIOTTGNQOVIVRKLDLSDTKSIRAMAKGFAEEKH 120  
 DB 61 LAQGARVYLACRVEKGEIVAKEIOTTGNQOVIVRKLDLSDTKSIRAMAKGFAEEKH 120  
 QY 121 LHVINNAGVMMCPYSKTADGPEMHIGVNLGHFLTHLLLEKESASPRIVNSSLAH 180  
 DB 121 LHVINNAGVMMCPYSKTADGPEMHIGVNLGHFLTHLLLEKESASPRIVNSSLAH 180  
 QY 181 HIGRIHFHNLQGEKFNAGLAICHSKLANILFTQELARLKSGGVTTSVHGTVQSELY 240  
 DB 181 HIGRIHFHNLQGEKFNAGLAICHSKLANILFTQELARLKSGGVTTSVHGTVQSELY 240  
 QY 241 RHSSFRMWMWLFSPFIKTPOQAGQVRLHCAITGELISGNHFSQCHVAVTSAQARNET 300  
 DB 241 RHSSFRMWMWLFSPFIKTPOQAGQVRLHCAITGELISGNHFSQCHVAVTSAQARNET 300  
 QY 301 IARLMDV 308  
 DB 301 IARLMDV 308

RESULT 10  
ABG96558  
ID ABG96558 standard; protein; 318 AA.  
XX  
AC ABG96558;  
XX  
DT 12-DEC-2002 (first entry)  
XX  
DE Human short chain dehydrogenase family member CGI-82.  
XX  
KW Short chain dehydrogenase; SDR; human; antimetabolic; pesticide;  
KW herbicide; DHR deficiency; phenylketonuria; galactosaemia III;  
KW dienoil CoA reductase deficiency; adrenal hyperplasia; ovarian cancer;  
KW adrenogenital syndrome; mineralocorticoid excess syndrome; breast cancer;  
KW male pseudohemaphroditism; Zellweger syndrome; Down's syndrome;  
KW polycystic kidney disease; Alzheimer's disease; retinitis pigmentosa;  
KW retinitis punctata albescens; arterial hypertension; follicular lymphoma;  
KW hepatocarcinogenesis; fungicide; antibiotic.  
XX  
OS Homo sapiens.  
XX  
FN W0200212544-A2.  
XX  
PD 14-FEB-2002.  
XX  
PF 07-AUG-2001; 2001WO-EP009140.  
XX  
PR 07-AUG-2000; 2000US-0223436P.  
XX  
PA (BION-) BIONETWORKS GMBH.  
XX  
PI Wilckens T;  
XX  
DR WPI; 2002-241770/29.  
XX  
PT Identifying or verifying members of the short chain dehydrogenase (SDR)  
PT family, useful for novel drug development (e.g. for the development of  
PT antimetabolic, pesticides or herbicides), by employing an algorithm using  
PT core SDR motifs.  
XX  
PS Disclosure; Fig 4; 168pp; English.  
XX  
CC The invention relates to identifying or verifying members of the short  
CC chain dehydrogenase (SDR) family comprises employing an algorithm using  
CC core SDR motifs (MT1-MT4 and MVL1-VL2 given in the specification) for  
CC searching members of the SDR family. Also included are a member of the  
CC SDR family identified with the method above, a method for providing  
CC modulators for members of the SDR family, a method for evaluation of lead  
CC -candidates for possible modulators of a member of the SDR family and a  
CC method for detecting clinically relevant polymorphisms or single  
CC nucleotide polymorphisms. The method is useful for screening SDR  
CC sequences and modulators of the SDR family. The method is especially  
CC useful as a platform for novel drug development. The SDRs can serve for  
CC the development of e.g. antimetabolic, pesticides or herbicides. The  
CC modulators may be especially useful for the propionylaxis, treatment of/and  
CC diagnosis of diseases (e.g. DHR deficiency, phenylketonuria, dienoil CoA  
CC reductase deficiency, galactosaemia III, adrenal hyperplasia,  
CC adrenogenital syndrome, mineralocorticoid excess syndrome, ovarian cancer,  
CC breast cancer, male pseudohemaphroditism, Zellweger syndrome, polycystic  
CC kidney disease, Alzheimer's disease, retinitis punctata albescens,  
CC retinitis pigmentosa, Down's syndrome, arterial hypertension, follicular  
CC lymphoma and hepatocarcinogenesis) particularly as a fungicide or  
CC antibiotic. The present sequence is one of 39 human SDR family members  
CC identified by the method of the invention  
XX  
SQ Sequence 318 AA;  
XX  
Query Match 95.1%; Score 1577; DB 5; Length 318;  
Best Local Similarity 98.7%; Pred. No. 5.9e-155;  
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MVELMFPULLLLPFLLYAAAPQIRKMLSSGVCSTVQLPKRVVVVGANTGIGETAKE 60

DB 1 MVELMFPULLLLPFLLYAAAPQIRKMLSSGVCSTVQLPKRVVVVGANTGIGETAKE 60  
QY 61 LAORGARVYIACRDVEKSELVAKKEIOTTGNQOVVRKLDJSDTKSIRAMAKGFABEKH 120  
DB 61 LAORGARVYIACRDVEKSELVAKKEIOTTGNQOVVRKLDJSDTKSIRAMAKGFABEKH 120  
QY 121 LHWIINNAGVMMCPYSKTADGEEMHIGVNHGHPLLTHLLLEKESAPSRIVNVSSLAH 180  
DB 121 LHWIINNAGVMMCPYSKTADGEEMHIGVNHGHPLLTHLLLEKESAPSRIVNVSSLAH 180  
QY 181 HGRHFHNLGCEKRYVMGLACVSKLANILFTQELARLKSGSVTTSVHPGTQCSBLV 240  
DB 181 HGRHFHNLGCEKRYVMGLACVSKLANILFTQELARLKSGSVTTSVHPGTQCSBLV 240  
QY 241 RHSSFMKMMWMLFSEFIKTPQCGAQTRLHCALTEGLEILSGNHPSDCHVAVWSAQARNET 300  
DB 241 RHSSFMKMMWMLFSEFIKTPQCGAQTSLHCALTEGLEILSGNHPSDCHVAVWSAQARNET 300  
QY 301 IARRLMDV 308  
DB 301 IARRLMDV 308  
RESULT 11  
ABB95234  
ID ABB95234 standard; protein; 318 AA.  
XX  
AC ABB95234;  
XX  
DT 19-JUL-2002 (first entry)  
XX  
DE Human P509S protein SEQ ID NO 339.  
XX  
KW Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant;  
KW gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN US2002022248-A1.  
XX  
PD 21-FEB-2002.  
XX  
PF 12-JAN-2001; 2001US-00759143.  
XX  
PR 25-FEB-1997; 97US-00806099.  
PR 01-AUG-1997; 97US-00904804.  
PR 10-FEB-1998; 98US-00020956.  
PR 25-FEB-1998; 98US-00030607.  
PR 14-JUL-1998; 98US-00115453.  
PR 23-SEP-1998; 98US-00159812.  
PR 15-JAN-1999; 99US-00232149.  
PR 09-APR-1999; 99US-00258946.  
PR 13-JUL-1999; 99US-00352616.  
PR 12-NOV-1999; 99US-00439313.  
PR 18-NOV-1999; 99US-00443686.  
PR 14-JAN-2000; 2000US-00483672.  
PR 27-MAR-2000; 2000US-00536857.  
PR 09-MAY-2000; 2000US-00568100.  
PR 12-MAY-2000; 2000US-00570737.  
PR 13-JUN-2000; 2000US-00593793.  
PR 27-JUN-2000; 2000US-00605783.  
PR 10-AUG-2000; 2000US-00636215.  
PR 29-AUG-2000; 2000US-00651236.  
PR 06-SEP-2000; 2000US-00652729.  
PR 02-OCT-2000; 2000US-00679426.  
PR 10-OCT-2000; 2000US-00685166.  
XX  
XX (XU0J/) XU J.  
PA (DILL/) DILLON D C.  
PA (MITC/) MITCHAM J L.  
PA (HARL/) HARLOCKER S L.  
PA (JIAN/) JIANG Y.

PA (KALO/) KALOS M. D.  
 PA (FANG/) FANGER G. R.  
 PA (RETT/) RETTER M. W.  
 PA (STOL/) STOLK J. A.  
 PA (DAYC/) DAY C. H.  
 PA (VEDV/) VEDVICK T. S.  
 PA (CART/) CARTER D.  
 PA (LISX/) LI S. X.  
 PA (WANG/) WANG A.  
 PA (SKEI/) SKEIKY Y. A. W.  
 PA (HEPL/) HEPLER W. T.  
 PA (HEND/) HENDERSON R. A.  
 XX  
 PI Xu J., Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;  
 XX  
 DR WPI: 2002-255649/30.  
 XX  
 PT New prostate-specific polynucleotides for diagnosing and treating  
 PT diseases, in particular prostate cancer, and as markers for the  
 PT progression of cancer.  
 PS  
 PS Claim 2; SEQ ID NO 339; 87bp; English.  
 XX  
 CC The present invention provides prostate-specific coding sequences and  
 CC their encoded proteins. These can be used in the diagnosis and treatment  
 CC of cancers, particularly prostate cancer. The present sequence is a  
 CC protein described in the invention  
 XX  
 SQ Sequence 318 AA;  
 XX  
 Query Match 95.1%; Score 1577; DB 5; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5.9e-155;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MVELMPFLLLLPFLLYMAAPQIRKMLSSGVCSTVQLPKVVVVTGANTGIGETAKE 60  
 DB 1 MVELMPFLLLLPFLLYMAAPQIRKMLSSGVCSTVQLPKVVVVTGANTGIGETAKE 60  
 QY 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNQOVLVRKLDLSDTKSIRAWAKGFABEKH 120  
 DB 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNQOVLVRKLDLSDTKSIRAFAGFLABEKH 120  
 QY 121 LHVAINNAGVMMCPYSKTADGFEMHIGVNLGHFLTLHLLEKLSAPSRIVNVSLLAH 180  
 DB 121 LHVAINNAGVMMCPYSKTADGFEMHIGVNLGHFLTLHLLEKLSAPSRIVNVSLLAH 180  
 QY 181 HLGRIHFHNLQGEKFTNAGLAYCHSKLANILFTQELARLKSGGVTTYSVHPGTVOSELY 240  
 DB 181 HLGRIHFHNLQGEKFTNAGLAYCHSKLANILFTQELARLKSGGVTTYSVHPGTVOSELY 240  
 QY 241 RHSSFMKMMWMLFSFFIKTPOGAQOTRLHCAITLGELEILSGNHFSDCHVAVWSAQARNET 300  
 DB 241 RHSSFMKMMWMLFSFFIKTPOGAQOTSLHCALTEGLEILSGNHFSDCHVAVWSAQARNET 300  
 QY 301 IARRLMDV 308  
 DB 301 IARRLMDV 308  
 XX  
 RESULT 12  
 ID ABR54346 standard; protein; 318 AA.  
 XX  
 AC ABR54346;  
 XX  
 DT 28-AUG-2003 (first entry)  
 XX  
 DE Prostate tumour specific protein sequence SEQ ID 339.  
 XX  
 KW Cytostatic; gene therapy; prostate-specific protein; PSP; human;  
 KW immune response; prostate cancer.

XX  
 OS Homo sapiens.  
 XX  
 PN WO200289747-A2.  
 XX  
 PD 14-NOV-2002.  
 XX  
 PF 09-MAY-2002; 2002WO-US014753.  
 XX  
 PR 09-MAY-2001; 2001US-00852911.  
 PR 29-JUN-2001; 2001US-00895814.  
 PR 10-DEC-2001; 2001US-00012896.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Xu J., Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Henderson RA;  
 PI Kalos MD, Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS;  
 PI Carter D, Li SX, Wang A, Skeiky YAW, Hepler WT, Hural J;  
 PI McNeill PD, Houghton RL, Vinals Y De Baasols, Foy TM, Watanabe Y;  
 PI Deng T;  
 XX  
 DR WPI: 2003-167130/16.  
 XX  
 PT New prostate-specific proteins and genes, useful in gene therapy,  
 PT particularly for stimulating an immune response in a patient, or treating  
 PT prostate cancer in a patient, as well as for diagnosing prostate cancer  
 PT in a patient.  
 PS  
 PS Example 1; Page 402-403; 691pp; English.  
 XX  
 CC The present invention relates to novel prostate-specific proteins (PSP)  
 CC and their coding sequences. The PSPs and their coding sequences are  
 CC useful for stimulating an immune response in a patient, or for treating  
 CC prostate cancer in a patient and for determining, detecting or diagnosing  
 CC the presence of a cancer in a patient. The present sequence was used to  
 CC illustrate the invention  
 XX  
 SQ Sequence 318 AA;  
 XX  
 Query Match 95.1%; Score 1577; DB 6; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5.9e-155;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MVELMPFLLLLPFLLYMAAPQIRKMLSSGVCSTVQLPKVVVVTGANTGIGETAKE 60  
 DB 1 MVELMPFLLLLPFLLYMAAPQIRKMLSSGVCSTVQLPKVVVVTGANTGIGETAKE 60  
 QY 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNQOVLVRKLDLSDTKSIRAWAKGFABEKH 120  
 DB 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNQOVLVRKLDLSDTKSIRAFAGFLABEKH 120  
 QY 121 LHVAINNAGVMMCPYSKTADGFEMHIGVNLGHFLTLHLLEKLSAPSRIVNVSLLAH 180  
 DB 121 LHVAINNAGVMMCPYSKTADGFEMHIGVNLGHFLTLHLLEKLSAPSRIVNVSLLAH 180  
 QY 181 HLGRIHFHNLQGEKFTNAGLAYCHSKLANILFTQELARLKSGGVTTYSVHPGTVOSELY 240  
 DB 181 HLGRIHFHNLQGEKFTNAGLAYCHSKLANILFTQELARLKSGGVTTYSVHPGTVOSELY 240  
 QY 241 RHSSFMKMMWMLFSFFIKTPOGAQOTRLHCAITLGELEILSGNHFSDCHVAVWSAQARNET 300  
 DB 241 RHSSFMKMMWMLFSFFIKTPOGAQOTSLHCALTEGLEILSGNHFSDCHVAVWSAQARNET 300  
 QY 301 IARRLMDV 308  
 DB 301 IARRLMDV 308  
 XX  
 RESULT 13  
 ID ADB75403  
 XX  
 AC ADB75403 standard; protein; 318 AA.  
 AC ADB75403;

XX 04-DEC-2003 (first entry)  
 DT Prostate cancer marker protein.  
 DE Prostate; cancer; cytostatic; gene therapy; marker.  
 KM Homo sapiens.  
 OS WO2003009614-A2.  
 PN 06-FEB-2003.  
 PD 25-JUL-2002; 2002WO-US023913.  
 PE 25-JUL-2001; 2001US-0307982P.  
 PR 22-AUG-2001; 2001US-0314356P.  
 PR 25-SEP-2001; 2001US-0325020P.  
 PR 12-DEC-2001; 2001US-0341746P.  
 PR 05-MAR-2002; 2002US-0362158P.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA Schlegel R, Monahan JE, Endege WO, Ganaravaru M, Gorbacheva B,  
 PI Hoerish S, Kamatkar S, Monsey AM, Glatt K, Zhao X, Anderson D,  
 XX WPI; 2003-24803/24.  
 DR New nucleic acid molecule, useful for diagnosing or treating prostate  
 PT cancer.  
 PT Claim 4; SEQ ID NO 227; 99pp; English.  
 PS The invention relates to newly discovered cancer markers associated with  
 XX the cancerous state of prostate cells. Also disclosed is a method of  
 CC assessing whether a patient is afflicted with prostate cancer. The method  
 CC of the invention involves assessing whether a patient is afflicted with  
 CC prostate cancer by comparing the level of expression of a marker in a  
 CC patient sample and the normal level of expression of the marker in the  
 CC control non-prostate cancer sample, where a significant increase in the  
 CC level of expression of the marker in the patient sample and the normal  
 CC level indicates that the patient is afflicted with prostate cancer.  
 CC Nucleic acids of the invention are useful for diagnosing or treating  
 CC prostate cancer, and may be useful in gene therapy. Sequences given in  
 CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 318 AA;  
 SQ  
 Query Match 95.1%; Score 1577; DB 7; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5,9e-155;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

DB 241 RHSSFFRMWMLFSFIKTPOQGAOTRLHCALTEGILEISGNHPSDCHVAVWSAQAARNET 300  
 QY 301 IARRLMDV 308  
 DB 301 IARRLMDV 308  
 RESULT 14  
 ADB13789  
 ID ADB13789 standard; protein; 318 AA.  
 XX ADB13789;  
 AC 18-DEC-2003 (first entry)  
 DT Human prostate specific protein P5095.  
 DE Human; prostate specific cDNA; cytostatic; immunostimulant; gene therapy;  
 XX cell therapy; vaccine; T-cell epitope;  
 KW class I major histocompatibility complex allele; MHC; prostate cancer;  
 KM tumour; antigen presenting cell.  
 OS Homo sapiens.  
 XX US2003185830-A1.  
 PN 02-OCT-2003.  
 PD 12-NOV-2002; 2002US-00294025.  
 PE 25-FEB-1997; 97US-00806099.  
 PR 01-AUG-1997; 97US-00904804.  
 PR 06-FEB-1998; 98US-00020955.  
 PR 25-FEB-1998; 98US-00030607.  
 PR 14-JUL-1998; 98US-00115453.  
 PR 23-SEP-1998; 98US-00159812.  
 PR 15-JAN-1999; 99US-00232149.  
 PR 09-APR-1999; 99US-00288946.  
 PR 13-JUL-1999; 99US-00352616.  
 PR 12-NOV-1999; 99US-00439313.  
 PR 18-NOV-1999; 99US-00443686.  
 PR 14-JAN-2000; 2000US-00483672.  
 PR 27-MAR-2000; 2000US-00536857.  
 PR 09-MAY-2000; 2000US-00568100.  
 PR 12-MAY-2000; 2000US-00570737.  
 PR 13-JUN-2000; 2000US-00593793.  
 PR 27-JUN-2000; 2000US-00605783.  
 PR 09-AUG-2000; 2000US-00636215.  
 PR 29-AUG-2000; 2000US-00651236.  
 PR 06-SEP-2000; 2000US-00657279.  
 PR 02-OCT-2000; 2000US-00679426.  
 PR 10-OCT-2000; 2000US-00685166.  
 PR 09-NOV-2000; 2000US-00709729.  
 PR 12-JAN-2001; 2001US-00759143.  
 PR 09-FEB-2001; 2001US-00780669.  
 PR 09-MAY-2001; 2001US-00852911.  
 PR 29-JUN-2001; 2001US-00895814.  
 PR 10-DEC-2001; 2001US-00012896.  
 PR 09-MAY-2002; 2002US-00144678.  
 XX (CORI-) CORIYA CORP.  
 PA Xu J, Stolk JA, Kalos MD;  
 PI WPI; 2003-756193/71.  
 DR N-PSDB: ADB13782.  
 XX New isolated polypeptide for use in a vaccine for stimulating an immune  
 PT response, or for treating or diagnosis cancer, preferably prostate  
 PT cancer.  
 XX Example 1; Page: 10pp; English.



XX Sequence 318 AA;

Query Match 95.1%; Score 1577; DB 7; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 5.9e-155;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY	1	MVELMPFLLLLLLPFLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGKETAKE	60
DB	1	MVELMPFLLLLLLPFLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGKETAKE	60
QY	61	LAQRGARVYLACRDVEKGEIVAKEIQTTGNOQVLYAKLDLSDTKSIRANAKGFKAEKX	120
DB	61	LAQRGARVYLACRDVEKGEIVAKEIQTTGNOQVLYAKLDLSDTKSIRAFKGFLEBKX	120
QY	121	LHWVINNAGVMMCPYSKTADGFENHIGVNHGHFLTHLLLEKLEKESAPSRIVNSSLAH	180
DB	121	LHVLINNAGVMMCPYSKTADGFENHIGVNHGHFLTHLLLEKLEKESAPSRIVNSSLAH	180
QY	181	HLGRIFHPNLOGEKEFYNAGLAYCHSKLANILFTOELARLKSGSVTTYSHPGTVOSELV	240
DB	181	HLGRIFHPNLOGEKEFYNAGLAYCHSKLANILFTOELARLKSGSVTTYSHPGTVOSELV	240
QY	241	RHSSFTMRMMWLFSFFIKTPOQGAQTRLHCLTGLEIILSGNHPSDCHVAMVSAQARNET	300
DB	241	RHSSFTMRMMWLFSFFIKTPOQGAQTSIHCLTGLEIILSGNHPSDCHVAMVSAQARNET	300
QY	301	IARRLMDV 308	
DB	301	IARRLMDV 308	

Search completed: October 13, 2004, 18:11:49  
 Job time : 161 secs

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 18:13:15 ; Search time 131 Seconds

(without alignments)  
778.380 Million cell updates/sec

Title: US-10-021-002-2

Perfect score: 1659 1 MWELMPELILLPLFYMA.....RNETTARRLMDVVTWQASQ 316

Sequence: 1 MWELMPELILLPLFYMA.....RNETTARRLMDVVTWQASQ 316

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1356558 seqs, 322682953 residues

Total number of hits satisfying chosen parameters: 1356558

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PTC\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PTCUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1659	100.0	316	13	US-10-021-002-2
2	1577	95.1	318	9	US-09-759-143-339
3	1577	95.1	318	9	US-09-780-669-339
4	1577	95.1	318	9	US-09-822-827-339
5	1577	95.1	318	9	US-09-895-793-339
6	1577	95.1	318	9	US-09-895-814-339
7	1577	95.1	318	13	US-10-012-896-339
8	1577	95.1	318	14	US-10-010-940-339
9	1577	95.1	318	14	US-10-205-823-227
10	1577	95.1	318	14	US-10-144-678A-339
11	1577	95.1	318	14	US-10-294-025-339
12	1577	95.1	318	15	US-10-080-334-98
13	1577	95.1	318	16	US-10-408-765A-3006
14	1577	95.1	333	14	US-10-131-410-147

15	1577	95.1	335	14	US-10-131-410-198	Sequence 198, App
16	1573	94.8	318	15	US-10-080-334-97	Sequence 97, App1
17	1347.5	81.2	316	15	US-10-080-334-99	Sequence 99, App1
18	1347.5	81.2	316	15	US-10-074-978A-424	Sequence 424, App
19	1345.5	81.1	353	15	US-10-074-978A-425	Sequence 425, App
20	1291	77.8	300	15	US-10-074-978A-422	Sequence 422, App
21	1279	77.1	293	15	US-10-074-978A-423	Sequence 423, App
22	1130	68.1	316	14	US-10-094-749-1760	Sequence 1760, App
23	1129	68.1	316	15	US-10-286-506-1	Sequence 1, App1
24	1059	63.8	316	15	US-10-080-334-95	Sequence 95, App1
25	1020	61.5	278	15	US-10-080-334-92	Sequence 92, App1
26	1005	60.6	304	15	US-10-080-334-96	Sequence 96, App1
27	806	48.6	246	14	US-10-168-274-18	Sequence 18, App1
28	705.5	42.5	331	10	US-09-946-374-116	Sequence 116, App
29	705.5	42.5	331	14	US-10-006-856A-116	Sequence 116, App
30	705.5	42.5	331	14	US-10-006-818A-116	Sequence 116, App
31	705.5	42.5	331	14	US-10-006-85A-116	Sequence 116, App
32	705.5	42.5	331	14	US-10-013-907A-116	Sequence 116, App
33	705.5	42.5	331	14	US-10-015-499A-116	Sequence 116, App
34	705.5	42.5	331	14	US-10-015-499A-116	Sequence 116, App
35	705.5	42.5	331	14	US-10-015-869A-116	Sequence 116, App
36	705.5	42.5	331	14	US-10-012-121A-116	Sequence 116, App
37	705.5	42.5	331	14	US-10-006-116A-116	Sequence 116, App
38	705.5	42.5	331	14	US-10-006-117A-116	Sequence 116, App
39	705.5	42.5	331	14	US-10-017-527A-116	Sequence 116, App
40	705.5	42.5	331	14	US-10-013-913A-116	Sequence 116, App
41	705.5	42.5	331	14	US-10-007-194A-116	Sequence 116, App
42	705.5	42.5	331	14	US-10-013-430A-116	Sequence 116, App
43	705.5	42.5	331	14	US-10-011-671A-116	Sequence 116, App
44	705.5	42.5	331	14	US-10-012-752A-116	Sequence 116, App
45	705.5	42.5	331	14	US-10-015-386A-116	Sequence 116, App

## ALIGNMENTS

RESULT 1  
US-10-021-002-2  
Sequence 2, Application US/10021002  
Publication No. US20020150578A1  
GENERAL INFORMATION:  
APPLICANT: He, et al.  
TITLE OF INVENTION: Human Prostatic Specific Reductase  
FILE REFERENCE: PFI50D2  
CURRENT APPLICATION NUMBER: US/10/021,002  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: PCT/US95/01827  
PRIOR FILING DATE: 1995-01-20  
PRIOR APPLICATION NUMBER: 08/464,400  
PRIOR FILING DATE: 1995-06-05  
PRIOR APPLICATION NUMBER: 09/123,386  
PRIOR FILING DATE: 1998-07-28  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 316  
TYPE: PRT  
ORGANISM: human  
US-10-021-002-2  
Query Match 100.0% Score 1659; DB 13; Length 316;  
Best Local Similarity 100.0% Pred No. 1.5e-151;  
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MWELMPELILLPLFYMAAPQIRKMLSSGVCTSTVQLPGKVVVVTGANTGIGETAKE 60  
1 MWELMPELILLPLFYMAAPQIRKMLSSGVCTSTVQLPGKVVVVTGANTGIGETAKE 60  
DB 1 MWELMPELILLPLFYMAAPQIRKMLSSGVCTSTVQLPGKVVVVTGANTGIGETAKE 60  
QY 61 LAQRGRVYIACDVKKGLVAKETOTTGNOOVLRKLDSDTSIRAMAGFAEEKH 120  
61 LAQRGRVYIACDVKKGLVAKETOTTGNOOVLRKLDSDTSIRAMAGFAEEKH 120  
DB 61 LAQRGRVYIACDVKKGLVAKETOTTGNOOVLRKLDSDTSIRAMAGFAEEKH 120  
QY 121 LHWINNAGVMOCPYSKTAADGPEMHIGVNLGHFLTLLEKLEKESAPSRIVNSSLAH 180  
121 LHWINNAGVMOCPYSKTAADGPEMHIGVNLGHFLTLLEKLEKESAPSRIVNSSLAH 180

```

Db      121 LHWINNAGVMMCPYSTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNSSLAH 180
Qy      181 HLGRIHFHNLQGEKFNAGIAYCHSKLANILFTQELARRLKSGVTTYVHPGTVOSELY 240
Db      181 HLGRIHFHNLQGEKFNAGIAYCHSKLANILFTQELARRLKSGVTTYVHPGTVOSELY 240
Qy      241 RHSSFRMWWWLSPFFIKTPOQGAQTRLHCALTEGLEILSGNHFSDCHVAMWSAQARNET 300
Db      241 RHSSFRMWWWLSPFFIKTPOQGAQTRLHCALTEGLEILSGNHFSDCHVAMWSAQARNET 300
Qy      301 IARRLMDVITCMASQ 316
Db      301 IARRLMDVITCMASQ 316

```

```

RESULT 2
US-09-759-143-339
; Sequence 339, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kaios, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stoik, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 339
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-759-143-339

```

```

Query Match      95.1%; Score 1577; DB 9; Length 318;
Best Local Similarity 98.7%; Pred. No. 1,2e-143;
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy      1 MVELMPPLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVTGANTGIGKETAKE 60
Db      1 MVELMPPLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVTGANTGIGKETAKE 60
Qy      61 LAORGARVYLACDVEKGEIVAKEIOTTGNQOVLVRKLDLSTKSIIRAMAKGFABEKG 120
Db      61 LAORGARVYLACDVEKGEIVAKEIOTTGNQOVLVRKLDLSTKSIIRAMAKGFABEKG 120
Qy      121 LHWINNAGVMMCPYSTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNSSLAH 180
Db      121 LHWINNAGVMMCPYSTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNSSLAH 180
Qy      181 HLGRIHFHNLQGEKFNAGIAYCHSKLANILFTQELARRLKSGVTTYVHPGTVOSELY 240
Db      181 HLGRIHFHNLQGEKFNAGIAYCHSKLANILFTQELARRLKSGVTTYVHPGTVOSELY 240
Qy      241 RHSSFRMWWWLSPFFIKTPOQGAQTRLHCALTEGLEILSGNHFSDCHVAMWSAQARNET 300

```

```

Db      241 RHSSFRMWWWLSPFFIKTPOQGAQTRLHCALTEGLEILSGNHFSDCHVAMWSAQARNET 300
Qy      301 IARRLMDV 308
Db      301 IARRLMDV 308

```

```

RESULT 3
US-09-780-669-339
; Sequence 339, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kaios, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stoik, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hurel, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780,669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 339
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-780-669-339

```

```

Query Match      95.1%; Score 1577; DB 9; Length 318;
Best Local Similarity 98.7%; Pred. No. 1,2e-143;
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy      1 MVELMPPLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVTGANTGIGKETAKE 60
Db      1 MVELMPPLLLLLPFLLYMAAPQIRKMLSSGVCTSTVQLPGKVVVVTGANTGIGKETAKE 60
Qy      61 LAORGARVYLACDVEKGEIVAKEIOTTGNQOVLVRKLDLSTKSIIRAMAKGFABEKG 120
Db      61 LAORGARVYLACDVEKGEIVAKEIOTTGNQOVLVRKLDLSTKSIIRAMAKGFABEKG 120
Qy      121 LHWINNAGVMMCPYSTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNSSLAH 180
Db      121 LHWINNAGVMMCPYSTADGFEMHIGVNLGHFLTHLLEKLSAPSRIVNSSLAH 180
Qy      181 HLGRIHFHNLQGEKFNAGIAYCHSKLANILFTQELARRLKSGVTTYVHPGTVOSELY 240
Db      181 HLGRIHFHNLQGEKFNAGIAYCHSKLANILFTQELARRLKSGVTTYVHPGTVOSELY 240
Qy      241 RHSSFRMWWWLSPFFIKTPOQGAQTRLHCALTEGLEILSGNHFSDCHVAMWSAQARNET 300
Db      241 RHSSFRMWWWLSPFFIKTPOQGAQTRLHCALTEGLEILSGNHFSDCHVAMWSAQARNET 300
Qy      301 IARRLMDV 308
Db      301 IARRLMDV 308

```



TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C26  
CURRENT APPLICATION NUMBER: US/09/895,614  
CURRENT FILING DATE: 2001-06-29  
NUMBER OF SEQ ID NOS: 990  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 339  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-895-814-339

Query Match 95.1%; Score 1577; DB 9; Length 318;  
Best Local Similarity 98.7%; Pred. No. 1,2e-143;  
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMFPDLLLLPFLLYMAAPQIRKMLSSGVCTSTVOLPGKVVVVGANTGIGKETAKE 60  
DB 1 MVELMFPDLLLLPFLLYMAAPQIRKMLSSGVCTSTVOLPGKVVVVGANTGIGKETAKE 60  
QY 61 LAQGARVYLAACRDVEKGEIVAKEIOTTGNOQVLVRKLDLSDTKSIRAWAKGFAEKH 120  
DB 61 LAQGARVYLAACRDVEKGEIVAKEIOTTGNOQVLVRKLDLSDTKSIRAFKGFLEEKH 120  
QY 121 LHWVNNAGVMMCPYSTADGFEMHIGVNHGFLTHLLEKLSKESAPSRIVNVSSLAH 180  
DB 121 LHWVNNAGVMMCPYSTADGFEMHIGVNHGFLTHLLEKLSKESAPSRIVNVSSLAH 180  
QY 181 HLGRIHFHNLQGEKFNAGLAICHSKLANILFTQELARLRKSGGVTTYSHPGTVOSELY 240  
DB 181 HLGRIHFHNLQGEKFNAGLAICHSKLANILFTQELARLRKSGGVTTYSHPGTVOSELY 240  
QY 241 RHSSFMKMMWMLFSFFIKTPQOGAQTSLHCALTEGLEILSGNHFSDCHVAMVSAQARNET 300  
DB 241 RHSSFMKMMWMLFSFFIKTPQOGAQTSLHCALTEGLEILSGNHFSDCHVAMVSAQARNET 300  
QY 301 IARRLMDV 308  
DB 301 IARRLMDV 308

RESULT 7  
US-10-012-896-339  
Sequence 339, Application US/10012896  
Publication No. US20020183251A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jianshun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yugui  
APPLICANT: Kalos, Michael D.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darlick  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skelky, Yasir A.W.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Hural, John  
APPLICANT: McNeill, Patricia D.  
APPLICANT: Houghton, Raymond L.  
APPLICANT: Vinals de Bassols, Carlotia  
APPLICANT: Roy, Teresa  
APPLICANT: Fanger, Gary R.  
APPLICANT: Wantanabe, Yoshihiro  
APPLICANT: Meagher, Madeleine Joy  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427C27

CURRENT APPLICATION NUMBER: US/10/012,896  
CURRENT FILING DATE: 2001-12-10  
NUMBER OF SEQ ID NOS: 1011  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 339  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-012-896-339

Query Match 95.1%; Score 1577; DB 13; Length 318;  
Best Local Similarity 98.7%; Pred. No. 1,2e-143;  
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMFPDLLLLPFLLYMAAPQIRKMLSSGVCTSTVOLPGKVVVVGANTGIGKETAKE 60  
DB 1 MVELMFPDLLLLPFLLYMAAPQIRKMLSSGVCTSTVOLPGKVVVVGANTGIGKETAKE 60  
QY 61 LAQGARVYLAACRDVEKGEIVAKEIOTTGNOQVLVRKLDLSDTKSIRAWAKGFAEKH 120  
DB 61 LAQGARVYLAACRDVEKGEIVAKEIOTTGNOQVLVRKLDLSDTKSIRAFKGFLEEKH 120  
QY 121 LHWVNNAGVMMCPYSTADGFEMHIGVNHGFLTHLLEKLSKESAPSRIVNVSSLAH 180  
DB 121 LHWVNNAGVMMCPYSTADGFEMHIGVNHGFLTHLLEKLSKESAPSRIVNVSSLAH 180  
QY 181 HLGRIHFHNLQGEKFNAGLAICHSKLANILFTQELARLRKSGGVTTYSHPGTVOSELY 240  
DB 181 HLGRIHFHNLQGEKFNAGLAICHSKLANILFTQELARLRKSGGVTTYSHPGTVOSELY 240  
QY 241 RHSSFMKMMWMLFSFFIKTPQOGAQTSLHCALTEGLEILSGNHFSDCHVAMVSAQARNET 300  
DB 241 RHSSFMKMMWMLFSFFIKTPQOGAQTSLHCALTEGLEILSGNHFSDCHVAMVSAQARNET 300  
QY 301 IARRLMDV 308  
DB 301 IARRLMDV 308

RESULT 8  
US-10-010-940-339  
Sequence 339, Application US/10010940  
Publication No. US20030088062A1  
GENERAL INFORMATION:  
APPLICANT: Xu, Jianshun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang, Yugui  
APPLICANT: Reed, Steven G.  
APPLICANT: Kalos, Michael  
APPLICANT: Fanger, Gary  
APPLICANT: Retter, Mark  
APPLICANT: Stolk, John  
APPLICANT: Day, Craig  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 210121.427D3  
CURRENT APPLICATION NUMBER: US/10/010,940  
CURRENT FILING DATE: 2001-12-05  
NUMBER OF SEQ ID NOS: 575  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 339  
LENGTH: 318  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-010-940-339

Query Match 95.1%; Score 1577; DB 14; Length 318;  
Best Local Similarity 98.7%; Pred. No. 1,2e-143;  
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMFPDLLLLPFLLYMAAPQIRKMLSSGVCTSTVOLPGKVVVVGANTGIGKETAKE 60

```

Db      1 MVELMPFLLLLLPFLLYMAAPDIRKMLSSGVCSTSVQCPGVVVTGANTGIGKETAKE 60
Qy      61 LAORGARVYLACRDVEKGEIVAKEIOTTGNQOVLRKLDLSDTKSIRAKAGFLAEK 120
Db      61 LAORGARVYLACRDVEKGEIVAKEIOTTGNQOVLRKLDLSDTKSIRAKAGFLAEK 120
Qy      121 LHWINNAVVMCPYSTADGFEHNGHILTHLLEKESAPSRIVNSSLAH 180
Db      121 LHWINNAVVMCPYSTADGFEHNGHILTHLLEKESAPSRIVNSSLAH 180
Qy      181 HLGRHFNHNGEKEFYNAAGLACHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELV 240
Db      181 HLGRHFNHNGEKEFYNAAGLACHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELV 240
Qy      241 RHSSFMWMMWLFSSFFIKTPQOGAOTRLHCALEGLISGNHPSDCHAVWVSAQAARNET 300
Db      241 RHSSFMWMMWLFSSFFIKTPQOGAOTRLHCALEGLISGNHPSDCHAVWVSAQAARNET 300
Qy      301 IARRLMDV 308
Db      301 IARRLMDV 308

```

RESULT 9

```

US-10-205-823-227
; Sequence 227, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatekar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 227
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-227

```

```

Query Match      95.1%; Score 1577; DB 14; Length 318;
Best Local Similarity 98.7%; Pred. No. 1,2e-143;
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      1 MVELMPFLLLLLPFLLYMAAPDIRKMLSSGVCSTSVQCPGVVVTGANTGIGKETAKE 60
Db      1 MVELMPFLLLLLPFLLYMAAPDIRKMLSSGVCSTSVQCPGVVVTGANTGIGKETAKE 60
Qy      61 LAORGARVYLACRDVEKGEIVAKEIOTTGNQOVLRKLDLSDTKSIRAKAGFLAEK 120

```

```

Db      61 LAORGARVYLACRDVEKGEIVAKEIOTTGNQOVLRKLDLSDTKSIRAKAGFLAEK 120
Qy      121 LHWINNAVVMCPYSTADGFEHNGHILTHLLEKESAPSRIVNSSLAH 180
Db      121 LHWINNAVVMCPYSTADGFEHNGHILTHLLEKESAPSRIVNSSLAH 180
Qy      181 HLGRHFNHNGEKEFYNAAGLACHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELV 240
Db      181 HLGRHFNHNGEKEFYNAAGLACHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELV 240
Qy      241 RHSSFMWMMWLFSSFFIKTPQOGAOTRLHCALEGLISGNHPSDCHAVWVSAQAARNET 300
Db      241 RHSSFMWMMWLFSSFFIKTPQOGAOTRLHCALEGLISGNHPSDCHAVWVSAQAARNET 300
Qy      301 IARRLMDV 308
Db      301 IARRLMDV 308

```

RESULT 10

```

US-10-144-678A-339
; Sequence 339, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Ranger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Hepler, William T.
; APPLICANT: Huxal, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals y de Baasols, Carlotca
; APPLICANT: Foy, Teresa W.
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Deng, Ta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427028
; CURRENT APPLICATION NUMBER: US/10/144,678A
; CURRENT FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 1033
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 339
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-678A-339

```

```

Query Match      95.1%; Score 1577; DB 14; Length 318;
Best Local Similarity 98.7%; Pred. No. 1,2e-143;
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy      1 MVELMPFLLLLLPFLLYMAAPDIRKMLSSGVCSTSVQCPGVVVTGANTGIGKETAKE 60
Db      1 MVELMPFLLLLLPFLLYMAAPDIRKMLSSGVCSTSVQCPGVVVTGANTGIGKETAKE 60
Qy      61 LAORGARVYLACRDVEKGEIVAKEIOTTGNQOVLRKLDLSDTKSIRAKAGFLAEK 120
Db      61 LAORGARVYLACRDVEKGEIVAKEIOTTGNQOVLRKLDLSDTKSIRAKAGFLAEK 120

```

```

QY 121 LHWINNAGVWMCPSYKTADGFEMHIGVNLGHFLTHLLLEKLESAPSRIVNVSIAH 180
DB 121 LHWLNNAGVWMCPSYKTADGFEMHIGVNLGHFLTHLLLEKLESAPSRIVNVSIAH 180
QY 181 HLGRIHFNLOGEKFYNAGLAICHSKLANILFTQELARRLKSGVTTYVHPGTVOSELY 240
DB 181 HLGRIHFNLOGEKFYNAGLAICHSKLANILFTQELARRLKSGVTTYVHPGTVOSELY 240
QY 241 RHSSFFMRMMWMLFSPFIKTPOGAGQTRJHCALTEGELISGNHPSDCHVAMVSAQARNET 300
DB 241 RHSSFFMRMMWMLFSPFIKTPOGAGQTRJHCALTEGELISGNHPSDCHVAMVSAQARNET 300
QY 301 IARRLMDV 308
DB 301 IARRLMDV 308

```

```

RESULT 11
US-10-294-025-339
; Sequence 339, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Kaios, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 339
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-025-339

```

```

Query Match 95.1%; Score 1577; DB 14; Length 318;
Best Local Similarity 98.7%; Pred. No. 1,2e-143;
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMPELILLPLFLYMAAPQIRKMLSSGVCTSTVQJPGKVVVYTGANTGIGETAKE 60
DB 1 MVELMPELILLPLFLYMAAPQIRKMLSSGVCTSTVQJPGKVVVYTGANTGIGETAKE 60
QY 61 LAORGARVYLACDVEKGEIVAEIOTTGNQGVLRKLDSDTKSIRAWAGFAAEKH 120
DB 61 LAORGARVYLACDVEKGEIVAEIOTTGNQGVLRKLDSDTKSIRAWAGFAAEKH 120
QY 121 LHWINNAGVWMCPSYKTADGFEMHIGVNLGHFLTHLLLEKLESAPSRIVNVSIAH 180
DB 121 LHWLNNAGVWMCPSYKTADGFEMHIGVNLGHFLTHLLLEKLESAPSRIVNVSIAH 180
QY 181 HLGRIHFNLOGEKFYNAGLAICHSKLANILFTQELARRLKSGVTTYVHPGTVOSELY 240
DB 181 HLGRIHFNLOGEKFYNAGLAICHSKLANILFTQELARRLKSGVTTYVHPGTVOSELY 240
QY 241 RHSSFFMRMMWMLFSPFIKTPOGAGQTRJHCALTEGELISGNHPSDCHVAMVSAQARNET 300
DB 241 RHSSFFMRMMWMLFSPFIKTPOGAGQTRJHCALTEGELISGNHPSDCHVAMVSAQARNET 300
QY 301 IARRLMDV 308
DB 301 IARRLMDV 308

```

```

RESULT 12
US-10-080-334-98
; Sequence 98, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shinkets, Richard A
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Kexuda, Ramesh
; APPLICANT: Splek, Kimberly A
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Malvanar, Uriel M
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Patcurajan, Meera
; APPLICANT: Gangolli, Bsha A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Liu, Xiaohong
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spaderna, Steven K
; APPLICANT: Zernusen, Bryan D
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-275
; CURRENT APPLICATION NUMBER: US/10/080,334
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,523
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/322,712
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/311,980
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/330,307
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/278,796
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/281,521
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/276,677
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,595
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/270,220
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/274,295
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/318,526
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/286,548
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/291,765
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/270,797
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/276,400
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/270,810
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-080-334-98

```

```

Query Match 95.1%; Score 1577; DB 15; Length 318;
Best Local Similarity 98.7%; Pred. No. 1,2e-143;
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMPELILLPLFLYMAAPQIRKMLSSGVCTSTVQJPGKVVVYTGANTGIGETAKE 60

```

Db 1 MVELMPELILLPLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60  
 QY 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNOQVLRKLDLSDTKSIRAMAKGFABEKH 120  
 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNOQVLRKLDLSDTKSIRAFKAGFLABEKH 120  
 QY 121 LHWYINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLESAPSRIVNVSILAH 180  
 121 LHWYINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLESAPSRIVNVSILAH 180  
 Db 181 HIGRIHFHNLQGEKFFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 181 HIGRIHFHNLQGEKFFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 QY 241 RHSSFMRMWMLFSPFIKTPOQGAQTRJHCAITBGLBILSGNHFSDCHVAVWSAQARNET 300  
 241 RHSSFMRMWMLFSPFIKTPOQGAQTRJHCAITBGLBILSGNHFSDCHVAVWSAQARNET 300  
 Db 301 IARRLMDV 308  
 301 IARRLMDV 308

## RESULT 13

US-10-408-765A-3006  
 ; Sequence 3006, Application US/10408765A  
 ; Publication No. US20040101874A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ghosh, Soumitra S.  
 ; APPLICANT: Fahy, Eoin D.  
 ; APPLICANT: Zhang, Bing  
 ; APPLICANT: Gibson, Bradford W.  
 ; APPLICANT: Taylor, Steven W.  
 ; APPLICANT: Glemm, Gary M.  
 ; APPLICANT: Warnock, Dale E.  
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
 ; FILE REFERENCE: 660088.465  
 ; CURRENT APPLICATION NUMBER: US/10/408,765A  
 ; CURRENT FILING DATE: 2003-04-04  
 ; NUMBER OF SEQ ID NOS: 3077  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3006  
 ; LENGTH: 318  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-408-765A-3006

Query Match 95.1%; Score 1577; DB 16; Length 318;  
 Best Local Similarity 98.7%; Pred. No. 1.2e-143;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MVELMPELILLPLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60  
 1 MVELMPELILLPLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60  
 Db 1 MVELMPELILLPLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60  
 1 MVELMPELILLPLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60  
 QY 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNOQVLRKLDLSDTKSIRAMAKGFABEKH 120  
 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNOQVLRKLDLSDTKSIRAFKAGFLABEKH 120  
 Db 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNOQVLRKLDLSDTKSIRAFKAGFLABEKH 120  
 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNOQVLRKLDLSDTKSIRAFKAGFLABEKH 120  
 QY 121 LHWYINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLESAPSRIVNVSILAH 180  
 121 LHWYINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLESAPSRIVNVSILAH 180  
 Db 121 LHWYINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLESAPSRIVNVSILAH 180  
 121 LHWYINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLESAPSRIVNVSILAH 180  
 QY 181 HIGRIHFHNLQGEKFFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 181 HIGRIHFHNLQGEKFFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 Db 181 HIGRIHFHNLQGEKFFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 181 HIGRIHFHNLQGEKFFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 QY 241 RHSSFMRMWMLFSPFIKTPOQGAQTRJHCAITBGLBILSGNHFSDCHVAVWSAQARNET 300  
 241 RHSSFMRMWMLFSPFIKTPOQGAQTRJHCAITBGLBILSGNHFSDCHVAVWSAQARNET 300  
 Db 241 RHSSFMRMWMLFSPFIKTPOQGAQTRJHCAITBGLBILSGNHFSDCHVAVWSAQARNET 300  
 241 RHSSFMRMWMLFSPFIKTPOQGAQTRJHCAITBGLBILSGNHFSDCHVAVWSAQARNET 300  
 QY 301 IARRLMDV 308  
 301 IARRLMDV 308

Db 301 IARRLMDV 308  
 301 IARRLMDV 308

## RESULT 14

US-10-131-410-147  
 ; Sequence 147, Application US/10131410  
 ; Publication No. US20030235915A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SPECHT, THOMAS  
 ; APPLICANT: HINZMANN, BERND  
 ; APPLICANT: SCHMITT, ARMIN  
 ; APPLICANT: PILASKY, CHRISTIAN  
 ; APPLICANT: DAHL, EDGAR  
 ; APPLICANT: ROSENTHAL, ANDRE  
 ; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST  
 ; TITLE OF INVENTION: TUMORS  
 ; FILE REFERENCE: SCH-1763  
 ; CURRENT APPLICATION NUMBER: US/10/131,410  
 ; CURRENT FILING DATE: 2002-04-25  
 ; PRIOR APPLICATION NUMBER: 09/646,673  
 ; PRIOR FILING DATE: 2000-09-20  
 ; PRIOR APPLICATION NUMBER: PCT/DE99/00908  
 ; PRIOR FILING DATE: 1999-03-19  
 ; NUMBER OF SEQ ID NOS: 202  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 147  
 ; LENGTH: 333  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-131-410-147

Query Match 95.1%; Score 1577; DB 14; Length 333;  
 Best Local Similarity 98.7%; Pred. No. 1.3e-143;  
 Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVELMPELILLPLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60  
 1 MVELMPELILLPLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60  
 Db 16 MVELMPELILLPLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 75  
 16 MVELMPELILLPLLYMAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 75  
 QY 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNOQVLRKLDLSDTKSIRAMAKGFABEKH 120  
 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNOQVLRKLDLSDTKSIRAFKAGFLABEKH 135  
 Db 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNOQVLRKLDLSDTKSIRAFKAGFLABEKH 135  
 61 LAORGARVYLACRDVEKGEIVAKEIOTTGNOQVLRKLDLSDTKSIRAFKAGFLABEKH 135  
 QY 121 LHWYINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLESAPSRIVNVSILAH 180  
 121 LHWYINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLESAPSRIVNVSILAH 180  
 Db 121 LHWYINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLESAPSRIVNVSILAH 195  
 121 LHWYINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLLEKLESAPSRIVNVSILAH 195  
 QY 181 HIGRIHFHNLQGEKFFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 181 HIGRIHFHNLQGEKFFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 240  
 Db 181 HIGRIHFHNLQGEKFFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 255  
 181 HIGRIHFHNLQGEKFFYNAGLAYCHSKLANILFTQELARRLKSGSVTTYSVHPGTVOSELY 255  
 QY 241 RHSSFMRMWMLFSPFIKTPOQGAQTRJHCAITBGLBILSGNHFSDCHVAVWSAQARNET 300  
 241 RHSSFMRMWMLFSPFIKTPOQGAQTRJHCAITBGLBILSGNHFSDCHVAVWSAQARNET 315  
 Db 241 RHSSFMRMWMLFSPFIKTPOQGAQTRJHCAITBGLBILSGNHFSDCHVAVWSAQARNET 315  
 241 RHSSFMRMWMLFSPFIKTPOQGAQTRJHCAITBGLBILSGNHFSDCHVAVWSAQARNET 315  
 QY 301 IARRLMDV 308  
 301 IARRLMDV 308  
 Db 301 IARRLMDV 323  
 301 IARRLMDV 323

## RESULT 15

US-10-131-410-198  
 ; Sequence 198, Application US/10131410  
 ; Publication No. US20030235915A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SPECHT, THOMAS  
 ; APPLICANT: HINZMANN, BERND  
 ; APPLICANT: SCHMITT, ARMIN  
 ; APPLICANT: PILASKY, CHRISTIAN  
 ; APPLICANT: DAHL, EDGAR  
 ; APPLICANT: ROSENTHAL, ANDRE  
 ; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST

TITLE OF INVENTION: TUMORS  
FILE REFERENCE: SCH-1763  
CURRENT APPLICATION NUMBER: US/10/131,410  
CURRENT FILING DATE: 2002-04-25  
PRIOR APPLICATION NUMBER: 09/646,673  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: PCT/DE99/00908  
PRIOR FILING DATE: 1999-03-19  
NUMBER OF SEQ ID NOS: 202  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 198  
LENGTH: 335  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-131-410-198

Query Match 95.1%; Score 1577; DB 14; Length 335;  
Best Local Similarity 98.7%; Pred. No. 1.3e-143;  
Matches 304; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 MVELMPELILLLPFLIYMAAPQIRKMLSSGVCSTVQLPGKYVVVTGANTGIGKETAKE 60
   |||||
Db 18 MVELMPELILLLPFLIYMAAPQIRKMLSSGVCSTVQLPGKYVVVTGANTGIGKETAKE 77
   |||||

QY 61 LAQGAAYVLACRDVEKGLVAKEIOTTGNQOVLRKLDLSDTKSIRAMAKGFKAEEKH 120
   |||||
Db 78 LAQGAAYVLACRDVEKGLVAKEIOTTGNQOVLRKLDLSDTKSIRAMAKGFKAEEKH 137
   |||||

QY 121 LHVINNAGVMMCPYSKTADGFEMHIGVNLGHFLTLHLLEKLEKESAPSRIYVSSLAH 180
   |||||
Db 138 LHVINNAGVMMCPYSKTADGFEMHIGVNLGHFLTLHLLEKLEKESAPSRIYVSSLAH 197
   |||||

QY 181 HLGRIHFHNLOGEKFYVAGLAYCHSKLANILFTQELARRLKSGVTTYSVHPGTVOSELY 240
   |||||
Db 198 HLGRIHFHNLOGEKFYVAGLAYCHSKLANILFTQELARRLKSGVTTYSVHPGTVOSELY 257
   |||||

QY 241 RHSSFMKMMWMLFSFPIKTPQGAQOTRLHCALTEGLEILSGNHFSDCHVAVWSAQARNET 300
   |||||
Db 258 RHSSFMKMMWMLFSFPIKTPQGAQOTSLHCALTEGLEILSGNHFSDCHVAVWSAQARNET 317
   |||||

QY 301 IARRIMDV 308
   |||||
Db 318 IARRIMDV 325
   |||||
```

Search completed: October 13, 2004, 18:24:53  
Job time: 133 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 18:05:04 ; Search time 40 Seconds

(without alignments)  
760.112 Million cell updates/sec

Title: US-10-021-002-2

Perfect score: 1659  
Sequence: 1 MWELMFPPLLILLPFLLYMA.....RNERTARRLMDVIVTCMASQ 316

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR 79: \*  
2: PIR: \*  
3: PIR: \*  
4: PIR: \*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486	29.3	350	2 T48275	hypothetical prote
2	476	28.7	304	2 T44727	probable oxidoredu
3	475.5	28.7	321	2 T02520	probable oxidoredu
4	470	28.3	303	2 E70848	probable oxidoredu
5	466.5	28.1	320	2 S42651	hypothetical prote
6	464.5	28.0	311	2 H70829	hypothetical prote
7	462.5	27.9	317	2 T10561	hypothetical prote
8	446	26.9	311	2 AD2021	hypothetical prote
9	440	26.5	316	2 D84206	probable oxidoredu
10	438	26.4	309	2 C83017	probable short cha
11	430.5	25.9	323	2 E87298	hypothetical prote
12	423.5	25.5	317	2 C70863	hypothetical prote
13	419.5	25.3	332	2 T13447	hypothetical prote
14	419.5	25.3	336	2 H75255	oxidoreductase, sh
15	406.5	24.5	339	2 T41570	hypothetical prote
16	400.5	24.1	333	2 T19314	hypothetical prote
17	389.5	23.5	400	2 C75365	dehydrogenase Atus
18	386.5	23.3	301	2 AE3195	dehydrogenase Atus
19	376.5	22.7	322	2 T04022	hypothetical prote
20	373.5	22.5	398	2 T33973	hypothetical prote
21	370.5	22.3	379	2 T23592	hypothetical prote
22	363	21.9	294	2 T05381	hypothetical prote
23	360.5	21.7	339	2 B86906	oxidoreductase yxd
24	349.5	21.1	329	2 T15910	hypothetical prote
25	349	21.0	204	2 T46363	hypothetical prote
26	342.5	20.6	925	2 T16235	hypothetical prote
27	341	20.6	311	2 T37155	probable oxidoredu
28	336	20.3	298	2 A47089	probable oxidoredu
29	323.5	19.5	297	2 S19842	probable oxidoredu

30	323.5	19.5	323	2 T29260	hypothetical prote
31	322	19.4	400	2 S20941	protochlorophyllid
32	321	19.3	321	2 S71468	protochlorophyllid
33	318.5	19.2	329	2 A12023	protochlorophyllid
34	314	18.9	315	2 T23597	hypothetical prote
35	314	18.9	398	2 J04146	protochlorophyllid
36	308	18.6	312	2 T17150	probable oxidoredu
37	307	18.5	388	2 S39394	protochlorophyllid
38	306	18.4	400	2 S30167	protochlorophyllid
39	305	18.4	388	2 S04783	protochlorophyllid
40	299	18.0	296	2 D82515	ketoreductase xfr27
41	299	18.0	401	2 T00897	protochlorophyllid
42	297.5	17.9	313	2 S08406	protochlorophyllid
43	297	17.9	401	2 T08936	protochlorophyllid
44	295.5	17.8	346	2 T51377	hypothetical prote
45	285.5	17.2	322	2 S76636	hypothetical prote

## ALIGNMENTS

RESULT 1  
T48275  
hypothetical protein T22P11.130 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T48275  
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, A.; Cross-References: UNIPROT:Q9L248; EMBL:AL162971  
A:Reference number: 224490  
A:Accession: T48275  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-350 <BEV>  
A:Cross-References: UNIPROT:Q9L248; EMBL:AL162971  
A:Experimental source: cultivar Columbia; BAC clone T22P11  
C:Genetics:  
A:Map position: 5  
A:Introns: 40/1; 136/2; 163/1; 214/2; 240/3; 269/1; 284/3  
A:Note: T22P11.130  
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match	Score	DB 2	Length	350
Best Local Similarity	37.1%	Pred. No. 5,8e-32		
Matches	114	Conservative	55	Mismatches 108; Indels 30; Gaps 5;
QY	34	TSVQLPKGVVVVVGANTGICETAKELAQGAPVYLACROVEKGLVAKELQTTGNOQ	93	
DB	26	TGIDATNLTAITGTTGIGETARVLKGAHVIGARNNGAENAKTEILRONANAR	85	
QY	94	VIVRKLDSDTKSIRAMAKGFVAEKHLHWI-----NNGVNMC	134	
DB	86	VTLQDLDSISIKAFREFHMLPLNLMYTSLSLIGFKAFAPFLNNGVWC	145	
QY	135	YKTKDGFPMHIGVNLGHLITLHLLKTKESAPS-----RIVVSSLAH---HLGITH	186	
DB	146	YQLSEDEIHLQFATNHIGHFLTNLLDTMKATATSGEGRILNWSVAHITYQEOIQ	205	
QY	187	FNLQEKFNAGLAVYCSKLANILFQELARLKGSGV---TTSVHGFTVOSELVRRSS	244	
DB	206	FQSIDICYSKRAVGSKLANIHNELSHQLOBEVNTITANSVHGLITLNFQHTA	265	
QY	245	FRKMMWMLSPFI-KTPOGAQTRLICAVTEGLEILSGNHSDCHVAVASQARNETAR	303	
DB	266	LIMRFLKFPSPFLMKNIPOGAATTCYVALHPSVKGVTKYFADCNVPSKLARDETIAQ	325	
QY	304	RLMDVIV 310		
DB	326	KLMDFSV 332		
RESULT 2	T44727			

Probable oxidoreductase [imported] - Mycobacterium leprae  
 C/Species: Mycobacterium leprae  
 C/Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
 C/Accession: T44727  
 R/James, K.D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.  
 Submitted to the EMBL Data Library, January 1998  
 A/Reference number: Z22831  
 A/Accession: T44727  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-304 <QMA>  
 A/Cross-references: UNIPROT:Q9ZBM5; EMBL:AL035159; PIDD:CAA22691.1  
 A/Experimental source: cosmid B1450  
 C/Genetics:  
 A/Note: MLCB1450.07  
 C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 28.7%; Score 476; DB 2; Length 304;  
 Best Local Similarity 41.0%; Pred. No. 3.2e-11;  
 Matches 120; Conservative 47; Mismatches 102; Indels 24; Gaps 9;

QY 35 STVQLP---GKVVYVGTANTGIGETAKELAQRGAVYLACRDVEKGEIVAKEIQTTCN 91  
 Db 5 TTADIPQGTGRVAVITGANTGLGYQTLALAEHGAIVLAVNLDKGDAAARITATSQ 64  
 QY 92 QQVIVRKLDLSDPKSIRAWAKGFKAEKHLHWINNAGVMMCPYSKTADGFEMHIGVNL 151  
 Db 65 NNVALGELDLASLESYRAAKKQARSDYDHDLDLNNAGVMTPKSTTKGFELQGTNHL 124  
 QY 152 GHFLTHLLLEKIKESAPSRIVNSSLAHL-GRIFHNLQGEKFNAGLAYCHSKLANI 210  
 Db 125 GHFAFGTGLDRLPLTVGSRVITVSSLSHRLFDIHFNDLOWECNNRVAAAGOSKLANI 184  
 QY 211 LFPQELARLRKSGGVTVTSV--HPTVQSELVHSSFMKMMWLFS---FPIKTPQGA 264  
 Db 185 LFTVEIQRR-ARQTTIAVAHAPGSSRELTIR-TLPILAFISVALFLTDAAATGA 241  
 QY 265 QTRHLCALTEGLEILSGNHFSD-----CHVAWSAQAARNETIAR-RLMDV 308  
 Db 242 LPTLRATADA--VLGGYFGPDGFAIRGHPKVVASNGSKSHVDROLRLMAV 292

RESULT 3  
 T02520  
 Probable oxidoreductase [imported] - Arabidopsis thaliana  
 N/Alternate names: hypothetical protein F13M22.4  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 09-Jul-2004  
 C/Accession: T02520; A84794  
 R/Rounseley, S.D.; Xauli, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, June 1998  
 A/Description: Arabidopsis thaliana chromosome II BAC F13M22 genomic sequence.  
 A/Reference number: Z14677  
 A/Accession: T02520  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-321 <ROU>  
 A/Cross-references: UNIPROT:O80924; EMBL:AC004684; NID:G3236234; PID:G3236237  
 A/Experimental source: cultivar Columbia  
 R/Lin, X.; Xauli, S.; Rounseley, S.D.; Shua, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Umayam, S.E.; Umayam, L.; Tallon, L.; eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salberg, S.L.; Fraser, C.M.; Venter, J. Nature 407, 761-768, 1999  
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A/Reference number: A84420; XUID:20083487; PMID:10617197  
 A/Accession: A84794  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-321 <STO>  
 A/Cross-references: GB:AE002093; NID:G3236237; PIDD:AA03625.1; GSPDB:GN00139  
 C/Genetics:  
 A/Gene: Atg37540; F13M22.4  
 A/Map position: 2

A/Introns: 40/1; 117/2; 144/1; 195/2; 221/3; 250/1.1.265/3  
 C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 28.7%; Score 475.5; DB 2; Length 321;  
 Best Local Similarity 38.9%; Pred. No. 3.7e-11;  
 Matches 112; Conservative 48; Mismatches 117; Indels 11; Gaps 4;

QY 34 TSTVQLPGKVVYVGTANTGIGETAKELAQRGAVYLACRDVEKGEIVAKEIQTTCN 93  
 Db 26 TQADASHLTAITIGTSGIGLEAARLARGAVLIIAANPRAESKEMILQNNPNAR 85  
 QY 94 VLVRKLDLSDPKSIRAWAKGFKAEKHLHWINNAGVMMCPYSKTADGFEMHIGVNL 153  
 Db 86 VDIQIDVSSIKSRSDVDQFLNVLNLTINNAGVMFPKFLTDGIESGATNIGH 145  
 QY 154 FLTHLLLEKIKESA---PSRIVNSSLAH--HLGRIFHNLQGEKFNAGLAYCHS 205  
 Db 146 FLTNLLDLDDKKSIRARSGVQGRIVNLSIAHTYVSEGIKFGINDPAGSERRAYGOS 205  
 QY 206 KLANILFTQELARLRKSGV--TTSVHPQTQSELVHSSFMKMMWLFSF-FIKTPQGA 262  
 Db 206 KLSNLSNLSNLSRRLQSEGVNITINSVPLVTTNLFRIYSGFSMKYFRAMTFIPLWKIIPQ 265  
 QY 263 GAQTRHLCALTEGLEILSGNHFSDCHVAWSAQAARNETIARRLMDV 310  
 Db 266 GAATTCVVALHPDLEGVTKGYFGDCNIVAPSKPATNNSLADKLMDPSV 313

RESULT 4  
 E70848  
 Probable oxidoreductase - Mycobacterium tuberculosis (strain H37RV)  
 C/Species: Mycobacterium tuberculosis  
 C/Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
 C/Accession: E70848  
 R/Coile, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Felkwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
 A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.  
 A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A/Reference number: A70500; MUID:98255987; PMID:9634230  
 A/Accession: E70848  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-303 <CO2>  
 A/Cross-references: UNIPROT:O53613; GB:AL021428; GB:AL123456; NID:G3261514; PIDD:CAA1624  
 A/Experimental source: strain H37RV  
 C/Genetics:  
 A/Gene: Rv0068  
 C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology <SADH>  
 F15-208/Domain: short-chain alcohol dehydrogenase homology

Query Match 28.3%; Score 470; DB 2; Length 303;  
 Best Local Similarity 40.6%; Pred. No. 9.7e-11;  
 Matches 113; Conservative 46; Mismatches 109; Indels 10; Gaps 5;

QY 41 GKVVYVGTANTGIGETAKELAQRGAVYLACRDVEKGEIVAKEIQTTCNQVLRKLD 100  
 Db 14 GRTAVITGANTGIGEFPAALAAHGAHVLAVALDKGQAARITEARPGAEVELELD 73  
 QY 101 LSPDKSIRAWAKGFKAEKHLHWINNAGVMMCPYSKTADGFEMHIGVNLGHFLTHLL 160  
 Db 74 LTIASLRARAAQAQKEDHQIDLDLNNAGVMYPRRTTADGEMQGTNHLGHFLTGL 133  
 QY 161 LEKIKESAPSRIVNSSLAHL-GRIFHNLQGEKFNAGLAYCHSKLANILFTQELAR 219  
 Db 134 IDRLPLVAGSGRVVTTISVGRIRPAALHFDPLQWERYRVAAYGAQKLANLFTVELQR 193  
 QY 220 LKSGGV-TTSVHPQTQSELVHSSFMKMMWLFSFIKTPQGAQTRLCQA---LLE 274  
 Db 194 LAPGTTIAVASHPSVNTVEVANNMPRLVAVAAIAPIMQDAELGALPTLRATDPAYR 253  
 QY 275 GLEILSGNHFSDCH---VAWSAQAARNETIARRLMDV 308



C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C/Accession: AD2021  
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yaeuda, M.; Tanaka, S.  
DNA Res. 8, 205-213, 2001  
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*  
A/Reference number: AB1807; MUID:21595285; PMID:11759840  
A/Accession: AD2021  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-311 <STO>  
A/Cross-references: UNIPROT:Q9HW93; GB:BA000019; PIDN:BAE73421.1; PID:G17130812; GSPDB:G  
A/Experimental source: strain PCC 7120  
C/Genetics:  
A/Gene: alr1722  
C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 26.8%; Score 446; DB 2; Length 311;  
Best Local Similarity 38.8%; Pred. No. 9e-29;  
Matches 108; Conservative 54; Mismatches 104; Indels 12; Gaps 5;

QY 41 GKVVVVGANTGIGKETAKELAQRGVYACRDEKSELVAKETOTTGNOQVLRKLD 100  
DB 16 GRLLVYSSSGIGYETARVANKQASVITAVNLDKNAKALIKQNKDADVKMLD 75  
QY 101 LSDTKIRAWAKGFKAEKELHVINNAGVMCPYSKTADGPEMTHGVNHLGHFLTHLL 160  
DB 76 LANLAVKKNFKNPKRYRLDILLNAGVMPYKTDGPELQGTNHLGHFLTHLL 135  
QY 161 LKIKESAPRIVVSSLAHLGRIFHNIOGE-KTYNAGLAYCHSKLANILFTQELAR 219  
DB 136 LEFLISTEGSRIVVSSGANNMKIFDDLNMBQRYAKWKAGDSKANILFTYELDKR 195  
QY 220 LKSGGVTTY--SVHPTVQSELVRHSSFMKMMWLFSPFKTPOQAGTFLCALTEGP- 276  
DB 196 LKNGIDITLVTAHPGTATLORTAGI--VYLVNGIYAQDITWALTLAAALEAGLX 253  
QY 277 -EILSGNHSDOCHVAVSAQ---ANETIARLMDV 308  
DB 254 GAHYFGPNFGEMRGYPIKESNELSKDQALAKKLMV 291

RESULT 9  
DB2026  
probable oxidoreductase [imported] - Halobacterium sp. NRC-1  
C/Species: Halobacterium sp. NRC-1  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
C/Accession: DB4206  
R/Ni, M.V.; Kennedy, S.P.; Mahafas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo, Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A/Authors: Hou, S.; Daniels, C.J.; Dennis, S.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
A/Title: Genome sequence of Halobacterium species NRC-1.  
A/Reference number: AB4160; MUID:20504483; PMID:11016950  
A/Accession: DB4206  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-316 <STO>  
A/Cross-references: UNIPROT:Q9HR22; GB:AE004437; NID:G10580084; PIDN:AA619016.1; GSPDB:G  
C/Genetics:  
A/Gene: yaJ01  
C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 26.5%; Score 440; DB 2; Length 316;  
Best Local Similarity 39.3%; Pred. No. 2.8e-28;  
Matches 119; Conservative 41; Mismatches 119; Indels 24; Gaps 7;

QY 28 LSSGVCSTVQLP---GKVVVVGANTGIGKETAKELAQRGVYACRDEKSELVAK 84  
DB 1 MSSDDMTAAALPFDQGRVVTGANSGLFETRAFAAGAHVAVMACSTRGEDARD 60  
QY 85 IOTTGNOQVLRKLDLSDTKIRAWAKGFKAEKELHVINNAGVMCPYSKTADGPEM 144

DB 61 IVALPGASLTVHEDLALDVAAPFADWTAEBSLHVANNAGVVAIRSETADQET 120  
QY 145 HIGVNLGHFLTHLLLEKKE-SAPRIVVSSLAHLGRIFHNIOGEKFNAGLAYC 203  
DB 121 QGVNHLGHVALTGLGLVLRRTSGETRVVTSQSGARRRIDFEDLOHAEVQKWEAYS 180  
QY 204 HSKLANILFTQELARRK--GSGVTYSVHPTVQSELVRH-----SFPKMMWLFSPF 256  
DB 181 OSKANILFAYELDRRLRAASASVTSVACHPGVATVQLRGQAQSRILAMRAANA 240  
QY 257 I--XTPOQAGTFLCALTEGELISGNHSDCHVAVV-----SAQARNETIARL 305  
DB 241 LVQSAEGAMPILLYATNPSID--GGEYIGPGVLMRGPBPQPSARSDEDIARL 298  
QY 306 MDV 308  
DB 299 MTV 301

RESULT 10  
C3017  
probable short chain dehydrogenase PA5031 [imported] - Pseudomonas aeruginosa (strain PA  
C/Species: Pseudomonas aeruginosa  
C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
C/Accession: C83017  
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardi, K.; Lim,  
N.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A/Reference number: AB2950; MUID:20437337; PMID:10984043  
A/Accession: C83017  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-309 <STO>  
A/Cross-references: UNIPROT:Q9HUB8; GB:AE004916; GB:AE004091; NID:G9951315; PIDN:AA60841  
A/Experimental source: strain PA01  
C/Genetics:  
A/Gene: PA5031  
C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 26.4%; Score 438; DB 2; Length 309;  
Best Local Similarity 38.8%; Pred. No. 4e-28;  
Matches 111; Conservative 44; Mismatches 113; Indels 18; Gaps 6;

QY 38 QLPKVVVVGANTGIGKETAKELAQRGVYACRDEKSELVAKETOTTGNOQVLR 97  
DB 11 QVGRLLVYSSSGIGYETARVANKQASVITAVNLDKNAKALIKQNKDADVKMLD 70  
QY 98 KLDLSDTKIRAWAKGFKAEKELHVINNAGVMCPYSKTADGPEMTHGVNHLGHFLTH 157  
DB 71 DLDLADLASIRACAGRQHRARLDLFFNAGVWFLPLRTRDGFEMQGTNHLGHFLTH 130  
QY 158 HLLKESAPRIVVSSLAHLGRIFHNIOGEKFNAGLAYCHSKLANILFTQELAR 217  
DB 131 GLILDSLLAPRRVVGMTSGFQFGLPLDDLNMBQRYAKWKAGDSKANILFTYELDKR 190  
QY 218 RLKSGGVTTYV--HPTVQSEL-----VRHSSFMKMMWLFSPFKTPOQAGTFL 268  
DB 191 RRAQQRVVLQSLAHLGHVATVNLQVAPAMSSSRGRVAMKANGAFQSAEMGALPAL 250  
QY 269 HCLTE---GLEILSGNHSDOCH---VAVSAQARNETIARLMDV 306  
DB 251 -SALTQQRWYGAVGPDRLVETRGVPAARIRRNARDGLAKRLM 295

RESULT 11  
BB7298  
hypothetical protein C03098 [imported] - Caulobacter crescentus  
C/Species: Caulobacter crescentus  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C/Accession: BB7298

R;Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Lab, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
N.; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapir, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: BE7298  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-323 <STO>  
A:Cross-references: UNIPROT:Q9AB36; GB:AE005673; NID:G13421559; PIDN:AAK22385.1; GSPDB:G  
C:Genetics:  
A:Gene: C00398  
C:Superfamily: ribitol dehydrogenase, short-chain alcohol dehydrogenase homology

Query Match 25.9%; Score 430.5; DB 2; Length 323;  
Best Local Similarity 38.6%; Pred. No. 1,7e-27;  
Matches 117; Conservative 35; Mismatches 110; Indels 41; Gaps 9;

39 LPKVVVVVTGANTGIGKETAKELAQRGARYVLACRDVEKELVAKETQTTGNOQVLRK 98  
24 LSGVVAIVTGAITGIVETALALAGAEVITARRKPELGEVANAINEBAGSKRVSPGM 83  
99 LLSSTDSIR---AWAKGFKAEEKHLVWVNNAGVMMCPYSTADGFEKHIGVNLGHF 154  
84 LLSSTDSIRHFFVWVG-----DRIDILINNAVVASPLMRTADGFEKHIGVNLGHF 137  
155 LITLLEKLKESKES-----PSRIYVSSLAHLGRIFHNLO-GEKRYNAGLAICSKLA 208  
138 LLSVLLPNIJAGKAGKRSRLVSLSSIGHRSRGIHFDPNYTRREKEMAYGQAKXA 197  
209 NILETQELARLKSGVTTYSVHPTQSELSVRSF--WMMWML-----PSFET 257  
198 NSLFAVGFDRFKQGHANAVMGGLTLPQRLHLEBGRALGMDENDQPREGF----- 253  
258 KTFQQAQTRHLCALTEGLEILSGNHSFCHVA-----W-VSAQARNETIARLMD 307  
254 KTEQGAATSWAAVSGSELEGVGLYLEDGQALPMGSKENPMTGVPHALDPEADRLMD 313  
308 VIV 310  
314 LSV 316

## RESULT 12

hypothetical protein Rv2263 - *Mycobacterium tuberculosis* (strain H37Rv)  
C:Species: *Mycobacterium tuberculosis*  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: C70863

R;Conor, R.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
Rajadream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.  
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: C70863  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-317 <COL>  
A:Cross-references: UNIPROT:O53537; GB:AL021925; GS:AL123456; NID:G2909556; PIDN:CAA1730  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: Rv2263  
C:Superfamily: ribitol dehydrogenase, short-chain alcohol dehydrogenase homology  
F:15-209/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 25.5%; Score 423.5; DB 2; Length 317;  
Best Local Similarity 35.4%; Pred. No. 6.3e-27;  
Matches 109; Conservative 59; Mismatches 109; Indels 31; Gaps 9;  
28 LSSGVCTSTVQLPKVVVVVTGANTGIGKETAKELAQRGARYVLACRDVEKELVAKETQTT 87

1 MAKDLVAIVPDSGLAITTANSGLFGRLRRISAGADIVYAIRNAKEAAVEERT 60  
88 TTGNOQVLRKLDSDTDSIRPAMAKGFKAEEKHLVWVNNAGVMMCPYS-XTADGFEKHI 146  
61 AVPDAKLITKMLDSSLSVAALGEQMAADRPIDLLINNGVMTTPRRVTTAOGFELQF 120  
147 GVNHLGHTLTHLLEKLKESAPSAIYVSSLAHLGRIFHNLO-GEKRYNAGLAICSK 206  
121 GSNHLGHFALTAHLPLPFAAQARVSLSSLAARRGIHFDLQFERSVAPMTAYGQSK 180  
207 LANILFTQELARLK--GSGVTTYSVHPTQSELSVRSF--WMMWML-----PSFET 253  
181 LAVMPAEELDRRSRAAGWGIISNAHGLTKNLQIAGPHGRDKPALMERLYTSRIF 240  
254 SFFI-KTFQQAQTRHLCALT--EG-----LEILSGNHSFCHVAVSAQARNET 300  
241 APFLMQEIEEGILPALYAAATPQADGAFYGPGRGREYVAGCG---VREAKVPAARRDA 296  
301 IARRLMDV 308  
297 DSKRLMEV 304

## RESULT 13

T13447

hypothetical protein T19F6.40 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: T13447; T13449

R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bandroft, I.; Mei  
submitted to the Protein Sequence Database, July 1999  
A:Reference number: 217587  
A:Accession: T13447

A:Molecule type: DNA  
A:Residues: 1-332 <BEV>  
A:Cross-references: UNIPROT:O22985; EMBL:AL109619; GSPDB:GN00062; ATSP:T19F6.40  
A:Experimental source: cultivar Columbia; BAC clone T19F6  
C:Genetics:  
A:Gene: ATSP:T19F6.40  
A:Map position: 4  
A:Intons: 41/1, 118/2, 145/1, 197/2, 224/3, 254/1, 269/3  
C:Superfamily: ribitol dehydrogenase, short-chain alcohol dehydrogenase homology

Query Match 25.3%; Score 419.5; DB 2; Length 332;  
Best Local Similarity 36.5%; Pred. No. 1.4e-26;  
Matches 109; Conservative 53; Mismatches 112; Indels 25; Gaps 6;

34 TSTVOLPKVVVVVTGANTGIGKETAKELAQRGARYVLACRDVEKELVAKETQTTGNOQ 93  
27 TENCRLRSITAVITATGIGETARVLAKGARLIFPARVAKAEKERVISEFPETE 86  
94 VVRLDLDSDTDSIRPAMAKGFKAEEKHLVWVNNAGVMMCPYSTADGFEKHIGVNLGH 153  
87 IVVMDLSSISASVNFADPESDLPLNLLINNAKLAHHAISDEIDMTFATNYIGH 146  
154 FLTLLEKLKESKES-----PSRIYVSS-----LAHLGRIFHNLO-GEKRYNAG 199  
147 FLTLNLLNKMIGTAEFTGVCGRIYVNSGIGHGFPSSDLIELRLISQPKCO---FDAT 202  
200 LAYGSKANILFTQELARLK--GSGVTTYSVHPTQSELSV--RHSFPMKMMWMLFSF 255  
203 RYVALSKANVHTLESRLQKIGAVNTVNCVHGAVRTPLTRDRBSGLTDLVFFLASK 262  
256 FKTPOQAQTRHLCALTEGLEILSGNHSFCHVAVSAQARNETIARLMD---DVIYT 311  
263 LKTVPQAATTCVYATNPRLVNVGSKYFTDCNETTPGCLGTNSSEATKMAASHILVT 321

## RESULT 14

H75255  
C:Species: *Deinococcus radiodurans*  
oxidoreductase, short-chain dehydrogenase/reductase family - *Deinococcus radiodurans* (s)

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C/Accession: H75255  
 R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A/RTitle: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A/Reference number: A75250; MIMD:20036896; PMID:10567266  
 A/Accession: H75255  
 A/Status: Preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-1336 <WHI>  
 A/Cross-references: UNIPROT:Q9R99; GB:AE002089; GB:AE000513; NID:g6460427; PIDD:AAFI213  
 A/Experimental source: strain R1  
 C/Genetics:  
 A/Map position: 1  
 C/Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 25.3%; Score 419.5; DB 2; Length 336;  
 Best Local Similarity 38.6%; Pred. No. 1,4e-26;  
 Matches 112; Conservative 32; Mismatches 125; Indels 21; Gaps 6;

QY 37 VDLPGKVVVTGANTGIGKETAKELAQGARYVLACRDYKGLYAKETITGNQOVLV 96  
 DB 39 VDLKGTAVVTGASGLGTETARALLAGHVILPVRDRAKGERVAELRQSTGTVELV 98  
 QY 97 RKLDSDTYSIRAMKGFABEKGHLVWVNNAGVMMCPYSKTADGFEMHIGVNLGHFL 156  
 DB 99 -DDLSGLASVRGAIEIROLAPRIHILINNAGVATPQSRVTDGFETGFGTNHGHFL 157  
 QY 157 THLLKLEKASPRIVNSSLAHNGRIHFNLQGEKF--YVAGLAYCHSKLANILFTQE 215  
 DB 158 TRELLPALMAAARVVALTSCHRRSDIVMDLNERPPYPMDAYGOSKTANALFVAG 217  
 QY 216 IARRLKSGVTTYVHPGVQSELYVHSSP--MRMMW-----LFSFKITPOQGAQRL 268  
 DB 218 LFORVADQGLTANAVHPGGIMTGLQKFEVPLEDQRMGODEHGTLNPFVKTPAEGASTSV 277  
 QY 269 HCLTREGLELNGNHSDC-HVAMVSAQAARNETI-----ARRLM 306  
 DB 278 WAATSPQLGVGGFLPLEDLOHSTPDESAFNPFGYKPYALDHESARRLM 327

## RESULT 15

T41570  
 hypothetical protein SPOC736.13 - fission yeast (Schizosaccharomyces pombe)  
 C/Species: Schizosaccharomyces pombe  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C/Accession: T41570  
 R/Mod, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.  
 submitted to the EMBL Data Library, May 1998  
 A/Reference number: 221991  
 A/Accession: T41570  
 A/Status: Preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-1339 <WCO>  
 A/Cross-references: UNIPROT:Q74959; EMBL:AL023705; PIDD:CAA19277.1; GSPDB:GN00068; SPDB:  
 A/Experimental source: strain 972H-; cosmid C736  
 C/Genetics:  
 A/Map position: 3  
 A/Map position: 3

Query Match 24.5%; Score 406.5; DB 2; Length 339;  
 Best Local Similarity 35.7%; Pred. No. 1.6e-25;  
 Matches 101; Conservative 60; Mismatches 107; Indels 15; Gaps 5;

QY 39 LRGKVVVVGANTGIGKETAKELAQGARYVLACRDYKGLYAKETITGNQOVLV 98  
 DB 40 LTRKVALVVGSSGGIGVTALETARAKAVYLAQRNEEKYQKVMQIHDEVHSHKIRFLR 99  
 QY 99 LDISDTYSIRAMKGFABEKGHLVWVNNAGVMMCPYSKTADGFEMHIGVNLGHFLTL 158

DB 100 LDIIDFESVYQAASFIKAEKHLVNNAGVNNPPELTKGVEIQITNYSHTLFT 159  
 QY 159 LLEKLEKESAPS-----RIVNSSLAH--HLGRHFNHLOGEKFNAGLA-YCHSKLA 208  
 DB 160 LLPTLRTAEBCRPGBVRIVHVASIAYLAQAPYSGIYFPDLNPHVILGTFAPYQOSKYA 219  
 QY 209 NILEFTOLARLKSGVTTYVHPGVQSELYVHSSPFMMWMMW-----LFSFKITPOQGAQ 265  
 DB 220 QILYSIALARLEKXYGISVSLHPGVIRLELTNYSFPFALKLEKSVFQITLADPTRGAM 279  
 QY 266 TRLHCALTEGL--EILSGNHFSDCHVAMVSAQAARNETIARRLM 306  
 DB 280 TSLVATSPKISKEHNLNGAYFTAAQRGILHRAHDPAFVEVELY 322

Search completed: October 13, 2004, 18:14:03  
 Job time : 53 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 13, 2004, 18:04:53 ; Search time 74 Seconds  
(without alignments)

2457.005 Million cell updates/sec

Title: US-10-021-002-2

Sequence: 1 MVELMPFLILLLPFLILYMA.....RNETIARLMDVIVTCMAQ 316

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: 1: UniProt\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1577	95.1	318	1 RDHB_HUMAN	Q8TC12 homo sapien
2	1577	95.1	318	2 CAG3461	Cag3461 homo sapi
3	1347.5	81.2	316	1 RDHB_MOUSE	Q9QYF1 m. retinol d
4	1291	77.8	300	2 Q9RLB8	Q9RLB8 mus musculu
5	1279	77.1	293	2 Q9RLB9	Q9RLB9 mus musculu
6	1259.5	75.9	407	2 Q6TUD3	Q6TUD3 mus musculu
7	1259.5	75.9	407	2 AAQ91067	AAQ91067 ratu no
8	1129	68.1	316	1 RDHC_HUMAN	Q96N18 homo sapien
9	1126	67.9	316	1 RDHC_BOVIN	P59837 bos taurus
10	1068	64.4	316	1 RDHC_MOUSE	Q6BYX4 mus musculu
11	961	57.9	319	2 Q6D678	Q6D678 brachydanto
12	832	50.2	188	2 Q9RLB5	Q9RLB5 mus musculu
13	799.5	48.2	291	2 Q6DCT1	Q6DCT1 brachydanto
14	762.5	46.0	331	2 Q6NY49	Q6NY49 brachydanto
15	747.5	45.1	336	2 AAH6739	AAH6739 brachydanto
16	747.5	45.1	336	2 Q7QJL8	Q7QJL8 anophelae 9
17	733.5	44.2	330	2 Q9V4Q3	Q9V4Q3 drosophila
18	733	44.2	325	2 Q8MZG9	Q8MZG9 drosophila
19	731.5	44.1	330	2 Q9V4Q2	Q9V4Q2 drosophila
20	728	43.9	300	2 Q9V4Q2	Q9V4Q2 drosophila
21	718	43.6	331	2 Q9V4P9	Q9V4P9 drosophila
22	714	43.3	334	2 Q7Q6B6	Q7Q6B6 anophelae 9
23	706.5	42.6	331	1 RDHD_HUMAN	Q6N187 homo sapien
24	705.5	42.5	331	2 Q6UX79	Q6UX79 homo sapien
25	705.5	42.5	331	2 AAQ88837	AAQ88837 homo sapi
26	707.5	42.3	334	1 RDHD_MOUSE	Q6CE87 mus musculu
27	699.5	42.2	329	2 Q6GPF0	Q6GPF0 xenopus lae
28	695.5	41.9	318	2 Q6P001	Q6P001 brachydanto
29	695.5	41.9	318	2 AAH65890	AAH65890 brachydanto
30	689	41.5	296	2 Q6MKN1	Q6MKN1 drosophila
31	681	41.0	318	2 Q6D6T2	Q6D6T2 brachydanto

32	680.5	41.0	314	2 Q6P3N2	Q6P3N2 xenopus tro
33	680.5	41.0	314	2 AAH63926	AAH63926 xenopus t
34	680	41.0	296	2 Q7QUT6	Q7QUT6 anophelae 9
35	672.5	40.5	406	2 Q9W404	Q9W404 drosophila
36	641.5	38.7	336	1 RDHE_HUMAN	Q9N185 homo sapien
37	638	38.5	334	1 RDHE_MOUSE	Q9E165 mus musculu
38	606.5	36.6	377	2 Q6UX07	Q6UX07 mus musculu
39	606.5	36.6	377	2 AAQ88929	AAQ88929 homo sapi
40	585	33.3	309	2 Q7T348	Q7T348 brachydanto
41	554.5	33.4	321	2 Q9LGI8	Q9LGI8 oryza sativ
42	543.5	32.8	327	2 Q6EBH7	Q6EBH7 homo sapien
43	539.5	32.5	260	2 Q96G88	Q96G88 homo sapien
44	535.5	32.3	288	2 Q7QHU4	Q7QHU4 anophelae 9
45	525.5	31.7	336	2 Q9VE80	Q9VE80 drosophila

## ALIGNMENTS

RESULT 1	ID	RDHB_HUMAN	STANDARD	PRT	318 AA.
AC	Q8TC12; Q9NRW0; Q9Y391;				
DT	10-OCT-2003 (Rel. 42; Created)				
DT	10-OCT-2003 (Rel. 42; Last sequence update)				
DT	05-JUL-2004 (Rel. 44; Last annotation update)				
DE	Retinol dehydrogenase 11 (EC 1.1.1.-) (Retinal reductase 1) (RALR1)				
DE	(Prostate short-chain dehydrogenase/reductase 1) (Androgen-regulated short-chain dehydrogenase/reductase 1) (HCV core-binding protein				
DE	HCBP12) (CGI-82).				
GN	Name=RDH11; Synonyms=PSDR1; ARSDR1;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Euteria; Primates; Carnivora; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	TISSUE=Prostate;				
RX	MEDLINE=21139101; PubMed=11245473;				
RA	Lin B., White J.T., Ferguson P.S., Wang S., Vessella R., Bumgarner R.,				
RA	True L.D., Hood L., Nelson P.S.,				
RT	"Prostate short-chain dehydrogenase reductase 1 (PSDR1): a new member				
RT	of the short-chain steroid dehydrogenase/reductase family highly				
RT	expressed in normal and neoplastic prostate epithelium."				
RT	Cancer Res. 61:1611-1618 (2001).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	TISSUE=Liver;				
RA	Li K., Wang L., Cheng J., Zhang L., Lu Y., Liu Y., Duan H.,				
RT	"Screening of HCV core binding protein from human liver cDNA library				
RT	by using yeast two hybrid system."				
RT	Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.				
RL	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RX	MEDLINE=20272150; PubMed=10810093;				
RA	Lai C.-H., Chou C.-Y., Chang L.-Y., Lin C.-S., Lin W.-C.,				
RT	"Identification of novel human genes evolutionarily conserved in				
RT	Caenorhabditis elegans by comparative proteomics."				
RT	Genome Res. 10:703-713 (2000).				
RN	[4]				
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).				
RC	TISSUE=Brain, Muscle, Placenta, and Prostate;				
RX	MEDLINE=22388257; PubMed=1477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.V., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Dischenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Browstein M.U., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.J.,				
RA	Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,				



RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fanhey U. Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Blitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Gittmow D.J., Schmitt U., Myers R.M.,  
 RA Botcheffeld Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnercher A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RL [5]  
 RP IDENTIFICATION AS A RETINAL REDUCTASE AND SUBCELLULAR LOCATION.  
 RX MEDLINE=22511122; PubMed=12036956; DOI=10.1074/jbc.M20288200;  
 RA Kestelivill N.Y., Chumakova O.V., Chezyrkina S.V., Belyaeva O.V.,  
 RA Lapshina E.A., Lin D.W., Matsunura M., Nelson P.S.,  
 RT "Evidence that the human gene for prostate short-chain  
 RT dehydrogenase/reductase (PSR1) encodes a novel retinal reductase  
 RT (RetR1)." ;  
 RL J. Biol. Chem. 277:28909-28915 (2002).  
 CC -1- FUNCTION: Exhibits an oxidoreductive catalytic activity towards  
 CC retinoids. Most efficient as an NADPH-dependent retinal reductase.  
 CC Displays high activity toward 9-cis and all-trans-retinol. Also  
 CC involved in the metabolism of short-chain aldehydes. No steroid  
 CC dehydrogenase activity detected.  
 CC -1- SUBCELLULAR LOCATION: Type-II membrane protein (Probable).  
 CC Associated with endoplasmic reticulum membrane.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q8RTC12-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q8RTC12-2; Sequence=VSP\_008159;  
 CC Note=No experimental confirmation available;  
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in the epithelial  
 CC cells of prostate, in both basal and luminal secretory cell  
 CC populations. Expressed at low levels in spleen, thymus, testis,  
 CC ovary, small intestine, colon, peripheral blood leukocytes,  
 CC kidney, adrenal gland and fetal liver. Not detected in prostatic  
 CC fibromuscular stromal cells, endothelial cells, or infiltrating  
 CC lymphocytes.  
 CC -1- INDUCTION: By androgens, in prostate cancer cells.  
 CC -1- PTM: Not glycosylated.  
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases  
 CC (SDR) family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/ebd-sib.ch/announce/>  
 CC or send an email to [license@ebi-sib.ch](mailto:license@ebi-sib.ch)).  
 CC -----  
 DR EMBL; AF167438; AAF89632.1; -;  
 DR EMBL; AF950508; AAK72049.1; -;  
 DR EMBL; AF951840; AAD30477.1; -;  
 DR EMBL; BC000112; AAH00112.1; -;  
 DR EMBL; BC011727; AAH1727.1; -;  
 DR EMBL; BC026274; AAH26274.1; -;  
 DR EMBL; BC037302; AAH37302.1; -;  
 DR EMBL; BC051291; AAH51291.1; -;  
 DR HSSP; Q28960; INSD.  
 DR HGENS; HGNC:17964; RDH11.  
 DR MIM; 607849; -;  
 DR GO; GO:0005622; C:intracellular; ISS.  
 DR GO; GO:0004745; F:retinol dehydrogenase activity; ISS.  
 DR GO; GO:0045494; P:photoreceptor maintenance; ISS.  
 DR GO; GO:0042572; P:retinol metabolism; ISS.  
 DR InterPro; IPR002198; Adh\_short.  
 DR InterPro; IPR002347; Adh\_short\_C2.  
 DR Pfam; PF00106; adh\_short\_1.  
 DR PRINTS; PRO0081; GDRDH.  
 DR PRINTS; PRO0080; SDRFAMILY.

DR		PROSITE PS00061; ADH SHORT; FALSE NEG.
KW		Alternative splicing; Endoplasmic reticulum; NADP; Oxidoreductase;
KV		Signal-anchor; Transmembrane.
FT	TRANSMEM	1 21 Signal-anchor for type II membrane protein (Potential).
FT	DOMAIN	22 318 Cytoplasmic (Potential).
FT	NP_BIND	48 54 NDPE (By similarity).
FT	ACT_SITE	202 202 By similarity.
FT	VARSPLIC	52 64 Missing (in isoform 2).
FT	CONFLICT	176 176 S -> F (in Ref. 4; AAHS1291).
FT	CONFLICT	294 294 A -> V (in Ref. 1).
FT	CONFLICT	316 316 P -> S (in Ref. 4; AAH26274).
SQ	SEQUENCE	318 AA; 35386 MM; SBQC36652774835 CRC64;
	Query Match	95.1%; Score 1577; DB 1; Length 318;
	Best Local Similarity	98.7%; Pred. No. 3, 5e-120;
	Matches 304; Conservative	1; Mismatches 3; Indels 0; Gaps
QY	1	MVELMFPILLILLPILLYAAPOIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60
DB	1	MVELMFPILLILLPILLYAAEQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60
QY	61	LARGARVYLACDVKSGELVAKEIOTTGNQOVLVRKLDSPTKSIFAMAKGFABEKH 12
DB	61	LARGARVYLACDVKSGELVAKEIOTTGNQOVLVRKLDSPTKSIFAFKGFABEKH 12
QY	121	LHWINNAGVMCPYSKTADGEFEMIIGHVHLGHFLTHLLLEKLKESAPRIIVNSSLAH 18
DB	121	LHWINNAGVMCPYSKTADGEFEMIIGHVHLGHFLTHLLLEKLKESAPRIIVNSSLAH 18
QY	181	HUGRIHFHNIOGKRFYNAGLAVCHSKLANILFTQGLARKLKSGVTTTTSVHPGTQSRLV 24
DB	181	HGRIRHFHNIOGKRFYNAGLAVCHSKLANILFTQGLARKLKSGVTTTTSVHPGTQSRLV 24
QY	241	RHSSFPRMMWWLPSPIKTPPOOGAQRHCALTTEGETELISGNHFDCHVAMVASQAARNET 30
DB	241	RHSSFPRMMWWLPSPIKITPPGOGAQTSLHCALTBEGTELISGNHFDCHVAMVASQAARNET 30
QY	301	IARRLMDV 308
DB	301	IARRLMDV 308
	RESULT 2	
CAG33461	ID	PRELIMINARY; FRT; 318 AA.
CAG33461	AC	CAG33461;
DT	01-JUN-2004	(TrEMBLrel. 27, Created)
DT	01-JUN-2004	(TrEMBLrel. 27, Last sequence update)
DT	01-JUN-2004	(TrEMBLrel. 27, Last annotation update)
DE		RDH11 protein.
GN		RDH11.
OS		Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP		SEQUENCE FROM N.A.
RA	EBert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;	
RT	"Cloning of human full open reading frames in Gateway(TM) system entry vector (pDONR201)."	
RL	Submitted (JUN-2004) to the EMBL/Genbank/DDBJ databases.	
DR	EMBL; CR457180; CAG33461.1; -	
SQ	SEQUENCE	318 AA; 35386 MM; SBQC36652774835 CRC64;
	Query Match	95.1%; Score 1577; DB 2; Length 318;
	Best Local Similarity	98.7%; Pred. No. 3, 5e-120;
	Matches 304; Conservative	1; Mismatches 3; Indels 0; Gaps
QY	1	MVELMFPILLILLPILLYAAPOIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60
DB	1	MVELMFPILLILLPILLYAAPQIRKMLSSGVCSTVQLPGKVVVVTGANTGIGETAKE 60



OY 61 LA0GARVYIACRVEKGEIYAKELIOTTGNOVYAKRLDSDTKSTRAMAKFKAEK 120  
 DB 61 LA0GARVYIACRVEKGEIYAKELIOTTGNOVYAKRLDSDTKSTRAMAKFKAEK 120  
 OY 121 LHWIINNAGVWMCPSYKTDGFEHNIHGHFLTHLLLEKLESAPSRIVVSLAH 180  
 DB 121 LHWIINNAGVWMCPSYKTDGFEHNIHGHFLTHLLLEKLESAPSRIVVSLAH 180  
 OY 181 HLGIHHRHNOGGEFYNAGLACHSKLANILFTOLARRLKSGSVTYSVHPTVQSELY 240  
 DB 181 HLGIHHRHNOGGEFYNAGLACHSKLANILFTOLARRLKSGSVTYSVHPTVQSELY 240  
 OY 241 RHSSFMWMMWLFPSFIKTPOGQACTRLHCALTEGLEIISGNHPSDCVAMVSAQARNET 300  
 DB 241 RHSSFMWMMWLFPSFIKTPOGQACTRLHCALTEGLEIISGNHPSDCVAMVSAQARNET 300  
 OY 301 IARRLMDV 308  
 DB 301 IARRLMDV 308  
 OY 301 IARRLMDV 308  
 DB 301 IARRLMDV 308  
 RESULT 3  
 RDHB MOUSE STANDARD; PRT; 316 AA.  
 AC Q9QYF1; Q9D0U5; (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DE 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Retinol dehydrogenase 11 (EC 1.1.1.-) (Retinal reductase 1) (Rald1)  
 DE (Prostate short-chain dehydrogenase/reductase 1) (Androgen-regulated  
 DE short-chain dehydrogenase/reductase 1) (Short-chain aldehyde  
 DE dehydrogenase) (SCALD) (Cell line WC/9.1L4 derived protein 1)  
 DE (M2C60).  
 GN Name=Rdh11; Synonym=Pedrl, Aredrl, Mdt1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC TISSUE=Testis;  
 RX MEDLINE=94289700; PubMed=8018917;  
 RA Hara T., Harada N., Mitsui H., Wura T., Ishizaka T., Miyajima A.;  
 RT "Characterization of cell phenotype by a novel cDNA library  
 RT interleukin-4-dependent cell line.";  
 RL Blood 84:189-199(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=2213316; PubMed=12137953; DOI=10.1016/S0378-1119(02)00718-7;  
 RA Moore S., Pritchard C., Lin B., Ferguson C., Nelson P.S.;  
 RT "Isolation and characterization of the murine prostate short-chain  
 RT dehydrogenase/reductase 1 (Pedrl) gene, a new member of the short-  
 RT chain steroid dehydrogenase/reductase family.";  
 RL Gene 293:149-160(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22803395; PubMed=12807874; DOI=10.1074/jbc.M304969200;  
 RA Kaush-Jacobi A., Ou J., Bashmakov Y.K., Shelton J.M., Richardson J.A.,  
 RA Goldstein J.L., Brown M.S.;  
 RT "Characterization of mouse short-chain aldehyde reductase (SCALD), an  
 RT enzyme regulated by steroid regulatory element-binding proteins.";  
 RL J. Biol. Chem. 278:32380-32385(2003).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Embryo;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Niyaido I., Otsu N., Saito K., Suzuki H., Yamanka I., Kiyosaka H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schombach C., Gojohori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Brad T., Brusio V., Chochia A., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frerz K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson J.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Kongaya A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A.,  
 RA Maglott D.R., Maltz L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numa K., Okido T., Pavan W.J., Petreia G.,  
 RA Petrovsky N., Pillai R., Pontius U.V., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sanderlin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavoian M., Zhu Y., Zimmer A., Carrinchi P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Komuro H., Nakamura M., Sakazume N., Sato K.,  
 RA Shitaki T., Waki K., Waki J., Waki J., Waki J., Waki J.,  
 RA Hara A., Hashizume W., Imocani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Breast tumor;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Berger J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Utshin T.B., Toshbnyk S., Carrinchi P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.O., Malek A.G., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Buterfield V.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Exhibits an oxidoreductive catalytic activity towards  
 CC retinoids. Most efficient as an NADPH-dependent all-trans-retinal  
 CC reductase. Also involved in the metabolism of short-chain  
 CC aldehydes.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).  
 CC Associated with endoplasmic reticulum membrane (By similarity).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=1;  
 CC Comment=A number of isoforms are produced.  
 CC Name=1;  
 CC -1- TISSUE SPECIFICITY: Expressed at higher level in liver and testis.  
 CC Expressed at lower levels in smooth muscle, thymus, submaxillary  
 CC gland and epididymis. In testis, expression is restricted to  
 CC pachytene spermatocytes. Also expressed in four layers of the  
 CC retina, including the outer segment of rods and cones.  
 CC -1- PTM: Not glycosylated (By similarity).  
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases  
 CC (SDR) family.  
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a  
 CC frameshift in position 311.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AB035959; BAA8521.1; ALT\_FRAME.  
 CC EMBL; AY039032; AAK91516.1; -  
 CC EMBL; AF474027; AAT79910.1; -  
 CC EMBL; AK004413; BAA23296.1; -  
 CC EMBL; BC018261; AAI18261.1; -  
 CC HSSP; Q28960; INSD.  
 CC MGD; MGI:102581; Rdn11.  
 CC GO; GO:0005622; C:intracellular; ISS.  
 CC GO; GO:0004745; F:retinol dehydrogenase activity; ISS.  
 CC GO; GO:0045494; P:photoreceptor maintenance; ISS.  
 CC GO; GO:0042572; P:retinol metabolism; ISS.  
 CC InterPro; IPR002196; ADH\_short.  
 CC InterPro; IPR002347; Adh\_short\_C2.  
 CC Pfam; PF00106; adh\_short\_1.  
 CC PRINTS; PR00081; GDRDH.  
 CC PRINTS; PR00080; SDRFAMILY.  
 CC PROSITE; PS00061; ADH\_SHORT.  
 CC KX Alternative splicing; Endoplasmic reticulum; NADP; Oxidoreductase;  
 CC Signal-anchor; Transmembrane.  
 CC TRANSMEM 1 21  
 CC FT DOMAIN 22 316  
 CC FT ACT\_SITE 199 199  
 CC FT NP\_BIND 45 51  
 CC FT CONFLICT 238 238  
 CC FT CONFLICT 279 279  
 CC SQ SEQUENCE 316 AA; 35148 MW; 62FAE25585CC05FE CRC64;

Query Match 81.2%; Score 1347.5; DB 1; Length 316;  
 Best local Similarity 84.3%; Pred. No. 1.8e-101;  
 Matches 257; Conservative 20; Mismatches 27; Indels 1; Gaps 1;

QY 5 MFPLLILL-LPFLLYMAAFOIRKMLSSGCVTSVQVLPGRVYVVTGANTGIGETAKELAQ 63  
 DB 1 MEFLLILLSPFLILYVTPKIRKMLSSGCVTSVQVLPGRVYVVTGANTGIGETAKELAQ 60  
 QY 64 RGRVYLACRDEKSELVAKEIOTTTGNOQVLRKDLSDPTSIRAMAGFPAEKEHLV 123  
 DB 61 RGRVYLACRDEKSELVAKEIOTTTGNOQVLRKDLSDPTSIRAMAGFPAEKEHLV 120  
 QY 124 WINNAGVWCPYSKTAADGFMHIGVNLGHFLTLHLLEKLESAPSRIVNVSSLAHLG 183  
 DB 121 LINNAGVWCPYSKTAADGFMHIGVNLGHFLTLHLLEKLESAPSRIVNVSSLAHLG 180  
 QY 184 RIFHNLOGKRFYNAGLAAGVSHKLANILFPGELARLKSGVTVVHGEGTQSELVRS 243  
 DB 181 RIFHNLOGKRFYNAGLAAGVSHKLANILFPGELARLKSGVTVVHGEGTQSELVRS 240  
 QY 244 SFRMWWMLPSPFIKTPOQGAQTRLCALTEGLEILSGNHFSDCHVAVWSAQRNETIAR 303  
 DB 241 SFRMWWMLPSPFIKTPOQGAQTRLCALTEGLEILSGNHFSDCHVAVWSAQRNETIAR 300  
 QY 304 RLMDV 308  
 DB 301 RLMDV 305

## RESULT 4

Q9R1R8 PRELIMINARY; PRT; 300 AA.  
 AC Q9R1R8; 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DE UBE-1b.  
 GN Name=Rdh11; Synonyms=Ube-1b;  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.  
 RC STRAIN=CD-1; TISSUE=Testis;  
 RA Goto M., Eddy E.M.;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases (SDR) family.  
 CC EMBL; AB030504; BAA82657.1; -  
 CC MGD; MGI:102581; Rdn11.  
 CC GO; GO:0016491; F:oxidoreductase activity; IEA.  
 CC GO; GO:0008152; P:metabolism; IEA.  
 CC InterPro; IPR002196; ADH\_short.  
 CC InterPro; IPR002347; Adh\_short\_C2.  
 CC Pfam; PF00106; adh\_short\_1.  
 CC PRINTS; PR00081; GDRDH.  
 CC PRINTS; PR00080; SDRFAMILY.  
 CC Oxidoreductase.  
 CC SEQUENCE 300 AA; 33240 MW; 7332230DB9057262 CRC64;

Query Match 77.8%; Score 1291; DB 2; Length 300;  
 Best local Similarity 85.6%; Pred. No. 6.7e-97;  
 Matches 244; Conservative 18; Mismatches 23; Indels 0; Gaps 0;

QY 24 IRKMLSSGCVTSVQVLPGRVYVVTGANTGIGETAKELAORGARVYLACRDEKSELVAK 83  
 DB 5 LRKMLSSGCVTSVQVLPGRVYVVTGANTGIGETAKELAORGARVYLACRDEKSELVAK 64  
 QY 84 EIOTTTGNOQVLRKDLSDPTSIRAMAGFPAEKEHLVWINNAGVWCPYSKTAADGFE 143  
 DB 65 EIOTTTGNOQVLRKDLSDPTSIRAMAGFPAEKEHLVWINNAGVWCPYSKTAADGFE 124  
 QY 144 MHIGVNLGHFLTLHLLEKLESAPSRIVNVSSLAHLGRIFHNLOGKRFYNAGLAVC 203  
 DB 125 MHIGVNLGHFLTLHLLEKLESAPSRIVNVSSLAHLGRIFHNLOGKRFYNAGLAVC 184  
 QY 204 HSKLANILFPGELARLKSGVTVVHGEGTQSELVRSFRRMWWMLPSPFIKTPOQ 263  
 DB 185 HSKLANILFPGELARLKSGVTVVHGEGTQSELVRSFRRMWWMLPSPFIKTPOQ 244  
 QY 264 AQTSLYCALTEGLEILSGNHFSDCHVAVWSAQRNETIARLMDV 308  
 DB 245 AQTSLYCALTEGLEILSGNHFSDCHVAVWSAQRNETIARLMDV 289

## RESULT 5

Q9R1R9 PRELIMINARY; PRT; 293 AA.

AC Q9R1R9; 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DE 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
 DE UBE-1a.  
 GN Name=Rdh11; Synonyms=Ube-1a;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.  
 RA Goto M., Eddy E.M.;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases (SDR) family.

CC EMBL; AB030503; BAA82656.1; -  
 CC MGD; MGI:102581; Rdn11.  
 CC GO; GO:0016491; F:oxidoreductase activity; IEA.  
 CC GO; GO:0008152; P:metabolism; IEA.  
 CC InterPro; IPR002196; ADH\_short.  
 CC InterPro; IPR002347; Adh\_short\_C2.  
 CC Pfam; PF00106; adh\_short\_1.

```
DR PRINTS: PRO0081; GDRDH.
DR PRINTS: PRO0080; SDRFAMILY.
KW Oxidoreductase.
SQ SEQUENCE 293 AA; 32442 MW; 1C4FFA9770F0E9D0 CRC64;

Query Match 77.1%; Score 1279; DB 2; Length 293;
Best Local Similarity 85.8%; Pred. No. 6.2e-96;
Matches 242; Conservative 17; Mismatches 23; Indels 0; Gaps 0;

QY 27 MLSSGVCSTVQLPGKVVVVGANTGIGKETAKELAGARVYLACRDVEKGEIVAKEIO 86
DB 1 MLSSGVCSTVQLPGKVVVVGANTGIGKETAKELAGARVYLACRDVEKGEIVAKEIO 60
QY 87 TTTGNQOVLVRKLDLSDTISRAMAKGPKAEKELHVMINNAVVMCPYSKTDGEMHI 146
DB 61 AVTGNQGVFVRKLDLADTKSIRAFKDFLAEEKHLHLINNAGVMCPYSKTDGEMHI 120
QY 147 GVNHLGHFLTHLLEKLEKESAPRIVNSSLAHHLGRHFHNLQSEKFFYNAGLAYCHSK 206
DB 121 GVNHLGHFLTHLLEKLEKESAPRIVNSSLGHHLGRHFHNLQSEKFFYNAGLAYCHSK 180
QY 207 LANILFTQELARLKSGSVTTYSVHPGTVOSELVRRHSFMRMMWMLFSPFIKTPOOGAQT 266
DB 181 LANILFTKELAKRLKSGSVTTYSVHPGTVSELRRYSIMRMWMLFVFPIKTPOEGAQT 240
QY 267 RLHCAITLEGLEISGNHPSDCHVAVWSAQAARNETIARRLMDV 308
DB 241 SLVCAVTEGLEISGSHFSDDQLAWWSQAAGNETIARRLMDV 282

RESULT 6
Q6TUD3 PRELIMINARY; PRT; 407 AA.
AC Q6TUD3;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE LRRGT00111.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Xu C.S., Chang C.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Yang K.J., Zhao L.F., Ma H., Wang L., Wang S.F., Xing X.K., Shen G.M.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
DR EMBL: AY387097; AAC91067.1; -
DR InterPro: IPR002198; Adh_short.
DR InterPro: IPR002347; Adh_short_C2.
DR Pfam: PF00106; adh_short; 1.
DR PRINTS: PRO0081; GDRDH.
DR PRINTS: PRO0080; SDRFAMILY.
KW Oxidoreductase.
SQ SEQUENCE 407 AA; 44779 MW; FB0FA5B23D7BAC4 CRC64;

Query Match 75.9%; Score 1259.5; DB 2; Length 407;
Best Local Similarity 71.2%; Pred. No. 3.5e-94;
Matches 245; Conservative 25; Mismatches 33; Indels 41; Gaps 2;
```

```
DB 166 IQATTGNSQVLRKLDLADTKSIRAFAGFLAEKELHILINNAGVMCPYSKTDGEM 225
QY 145 HGVNHLGHFLTHLLEKLEKESAPRIVNSSLAHHLGRHFHNLQSEKFFYNAGLAYCH 204
DB 226 HGVNHLGHFLTHLLEKLEKESAPRIVNSSLAHHLGRHFHNLQSEKFFYNAGLAYCH 285
QY 205 SKLANILFTQELARLKSGSVTTYSVHPGTVOSELVRRHSFMRMMWMLFSPFIKTPOOGA 264
DB 286 SKLANILFTKELARLKSGSVTTYSVHPGTVSELRRYSIMRMWMLFVFPIKTPOOGA 345
QY 265 QTRHCAITLEGLEISGNHPSDCHVAVWSAQAARNETIARRLMDV 308
DB 346 QTRHCAVTEGLEISGSHFSDDQLAWWSQAAGNETIARRLMDV 389

RESULT 7
AAQ91067 PRELIMINARY; PRT; 407 AA.
AC AAQ91067;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE LRRGT00111.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Xu C.S., Chang C.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Yang K.J., Zhao L.F., Ma H., Wang L., Wang S.F., Xing X.K., Shen G.M.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL "Liver regeneration after PH.";
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY387097; AAC91067.1; -
SQ SEQUENCE 407 AA; 44779 MW; FB0FA5B23D7BAC4 CRC64;

Query Match 75.9%; Score 1259.5; DB 2; Length 407;
Best Local Similarity 71.2%; Pred. No. 3.5e-94;
Matches 245; Conservative 25; Mismatches 33; Indels 41; Gaps 2;
```

```
QY 6 FPLLLLLPFL---LLYMAAPQIRKMLSSGVCSTVQLPGKVVVVGANTGIGKETAKELA 62
DB 46 FCLFLFLAFLPISSTIYFLTPGSRKMLSCGVCISNVQLSGKVAIVTGANTGIGKETAKDLA 105
QY 63 QR-----GARYYLAACRDVEKGEIVAKE 84
DB 106 RRGKCTCSQLTSADVSPLTIVMVLPRLGIDWPFPRFWFGARVYLACRDWQKGEIVASE 165
QY 85 IOTTGNOQVLRKLDLSDTISRAMAKGPKAEKELHVMINNAVVMCPYSKTDGEM 144
```

```
QY 6 FPLLLLLPFL---LLYMAAPQIRKMLSSGVCSTVQLPGKVVVVGANTGIGKETAKELA 62
DB 46 FCLFLFLAFLPISSTIYFLTPGSRKMLSCGVCISNVQLSGKVAIVTGANTGIGKETAKDLA 105
QY 63 QR-----GARYYLAACRDVEKGEIVAKE 84
DB 106 RRGKCTCSQLTSADVSPLTIVMVLPRLGIDWPFPRFWFGARVYLACRDWQKGEIVASE 165
QY 85 IOTTGNOQVLRKLDLSDTISRAMAKGPKAEKELHVMINNAVVMCPYSKTDGEM 144
DB 166 IQATTGNSQVLRKLDLADTKSIRAFAGFLAEKELHILINNAGVMCPYSKTDGEM 225
QY 145 HGVNHLGHFLTHLLEKLEKESAPRIVNSSLAHHLGRHFHNLQSEKFFYNAGLAYCH 204
DB 226 HGVNHLGHFLTHLLEKLEKESAPRIVNSSLAHHLGRHFHNLQSEKFFYNAGLAYCH 285
QY 205 SKLANILFTQELARLKSGSVTTYSVHPGTVOSELVRRHSFMRMMWMLFSPFIKTPOOGA 264
DB 286 SKLANILFTKELARLKSGSVTTYSVHPGTVSELRRYSIMRMWMLFVFPIKTPOOGA 345
QY 265 QTRHCAITLEGLEISGNHPSDCHVAVWSAQAARNETIARRLMDV 308
DB 346 QTRHCAVTEGLEISGSHFSDDQLAWWSQAAGNETIARRLMDV 389
```



```

ID  RDHC_BOVIN  STANDARD;  PRT;  316 AA.
AC  P59837.
DT  10-OCT-2003 (Rel. 42, Created)
DT  10-OCT-2003 (Rel. 42, Last sequence update)
DE  05-JUL-2004 (Rel. 44, Last annotation update)
DE  Retinol dehydrogenase 12 (EC 1.1.1.-) (Double substrate-specificity
DE  short chain dehydrogenase/reductase 2).
GN  Name=RDH12; Synonyms=DSSDR2;
OS  Bos taurus (Bovine); Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Eukaryota; Metazoa; Chordata; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC  Bovinae; Bos.
OC  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OX  NCBI_TaxID=9913;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Retina;
RX  MEDLINE=22323226; PubMed=12226107; DOI=10.1074/jbc.M208892200;
RA  Haeseleer F., Jang G.-F., Imanishi Y., Driesen C.A.G.G.,
RA  Macsumura M., Nelson P.S., Palczewski K.;
RT  "Dual-substrate specificity short chain retinol dehydrogenases from
RT  the vertebrate retina";
RL  J. Biol. Chem. 277:45537-45546(2002).
CC  -1- FUNCTION: Exhibits an oxidoreductive catalytic activity towards
CC  retinoids. Most efficient as an NADPH-dependent retinal reductase.
CC  Displays high activity toward 9-cis and all-trans-retinol. Also
CC  involved in the metabolism of short-chain aldehydes. No steroid
CC  dehydrogenase activity detected. Might be the key enzyme in the
CC  formation of 11-cis-retinol from 11-cis-retinol during
CC  regeneration of the cone visual pigments (By similarity).
CC  -1- TISSUE SPECIFICITY: Expressed in the eyes.
CC  -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC  (SDR) family.
CC  -----
CC  This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.emb-ch/announce/
CC  or send an email to license@eb-sib.ch).
CC  -----
EMBL: AY115489; AAM51556.1; -
DR  GO: GO:0005622; C:intracellular; ISS.
DR  GO: GO:0004745; P:retinol dehydrogenase activity; ISS.
DR  GO: GO:0045494; P:photoreceptor maintenance; ISS.
DR  GO: GO:0042572; P:retinol metabolism; ISS.
DR  InterPro: IPR002199; ADH_short.
DR  InterPro: IPR002347; Adh_short_C2.
DR  Pfam: PF01065; adh_short; 1.
DR  PRINTS: PR00081; GDRDH.
DR  PRINTS: PR00080; SDRFAMILY.
DR  PROSITE: PS00061; ADH_SHORT; FALSE_NEG.
KW  NADP; Oxidoreductase; Vision.
FT  NADP (By similarity).
FT  ACT_SITE 200 200 By similarity.
SQ  SEQUENCE 316 AA; 35171 MW; 258263092CA4185 CRC64;
Query Match 67.9%; Score 1126; DB 1; Length 316;
Best Local Similarity 70.9%; Pred. No. 2e-83;
Matches 217; Conservative 36; Mismatches 51; Indels 2; Gaps 1;

```

```

QY  193 GRIFHNLOGEFYVNAGLAYCHSKLANTLFTQELARLKGSGVTYVHPGTVOSELVRH 242
DB  181 GKIRFHDLDGDYRYVNLGFAVYCHSKLANVLPRELKRLKGVTVVYAAHPGVRSKLVHR 240
QY  243 SEFMFWMMWLFSEFFIKTPOQGAOTRLHCLTEGLEILSGNHSPDCHVAVWSAQNREFTA 302
DB  241 SFLICLWRLFSPLKTTWEGAGQSLHCLAEGLPEPSGKYSDDCKTWVSPRANNNKTA 300
QY  303 RLRLMDV 308
DB  301 ERLMTV 306
RESULT 10
ID  RDHC_MOUSE  STANDARD;  PRT;  316 AA.
AC  Q8BYK4; Q91WA5; Q9D1Y4;
DT  10-OCT-2003 (Rel. 42, Created)
DT  10-OCT-2003 (Rel. 42, Last sequence update)
DE  05-JUL-2004 (Rel. 44, Last annotation update)
DE  Retinol dehydrogenase 12 (EC 1.1.1.-).
GN  Name=RDH12;
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=spinal cord;
RX  MEDLINE=22354683; PubMed=1246651; DOI=10.1038/nature01266;
RA  Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA  Mikado I., Osato N., Saito R., Suzuki H., Yamanaoka I., Kiyosawa H.,
RA  Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,
RA  Baldarelli R., Hill D.P., Bull C., Hume D.A., Quackenbush J.,
RA  Schramm L.M., Kanpin A., Matsuda H., Batalov S., Beisel K.W.,
RA  Blake U.A., Brad D., Bruste V., Chochia C., Corbani D.E., Cousins S.,
RA  Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA  Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA  Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA  Karai A., Kawai H., Kawasawa Y., Kedzierski R.W., King B.L.,
RA  Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA  Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA  Nagasawa T., Nunata K., Okido T., Pavan W.O., Petrea G., Pesole G.,
RA  Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA  Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA  Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA  Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA  Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA  Wilming L.G., Wymshav-Boris A., Yanagisawa M., Yang I., Yang L.,
RA  Yuan Z., Zavoian M., Zhu Y., Zimmer A., Carrinci P., Hayatsu N.,
RA  Hirozane-Kishikawa T., Konno H., Nakamura M., Sakakume N., Sato K.,
RA  Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,
RA  Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA  Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA  Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA  Birney E., Hayashizaki Y.,
RT  Analysis of the mouse transcriptome based on functional annotation of
RT  60,770 full-length cDNAs.
RL  Nature 420:563-573(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Retina;
RX  MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,
RA  Altshuler S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.U., Uscid T.B., Toshiyuki S., Carrinci P., Prange C.,
RA  Raha S.S., Loggellano N.A., Peters J.O., Abramson R.D., Mullany S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.D., Malek J.A., Guinard P.H.,
RA  Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,

```

RA Villalon D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman A.C., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -1- FUNCTION: Exhibits an oxidoreductive catalytic activity towards  
 CC retinoids. Most efficient as an NADPH-dependent retinal reductase.  
 CC Displays high activity toward 9-cis and all-trans-retinol. Also  
 CC involved in the metabolism of short-chain aldehydes. No steroid  
 CC dehydrogenase activity detected. Might be the key enzyme in the  
 CC formation of 11-cis-retinal from 11-cis-retinol during  
 CC regeneration of the cone visual pigments (By similarity).  
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases  
 CC (SDR) family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AK020927; BAB32258.1; -;  
 DR EMBL: AK039233; BAC30288.1; -;  
 DR EMBL: BC016204; AAL16204.1; -;  
 DR HSSP: Q29860; 1NSD.  
 DR MGI: 1925224; Rdh12.  
 DR GO: GO:0005622; C:intracellular; ISS.  
 DR GO: GO:0004745; P:retinol dehydrogenase activity; ISS.  
 DR GO: GO:0045494; P:photoreceptor maintenance; ISS.  
 DR GO: GO:0042572; P:retinol metabolism; ISS.  
 DR InterPro: IPR002198; ADH\_short.  
 DR InterPro: IPR002347; Adh\_short\_C2.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDRDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 DR PROSITE: PS00061; ADH\_SHORT; FALSE\_NEG.  
 KM NADP: Oxidoreductase: Vision.  
 FT NP\_BIND 46 52 NADP (By similarity).  
 FT ACT\_SITE 200 200 By similarity.  
 FT CONFLICT 114 125 Missing (in Ref. 2).  
 FT CONFLICT 120 120 H -> D (in Ref. 1; BAB32258).  
 SQ SEQUENCE 316 AA; 35292 MW; CF5745B6710A6148 CRC64;  
 Query Match 64.4%; Score 1068; DB 1; Length 316;  
 Best Local Similarity 67.6%; Pred. No. 1,1e-78;  
 Matches 207; Conservative 41; Mismatches 56; Indels 2; Gaps 1;

QY 303 RRLMDV 308  
 |||||  
 DB 301 ERLMNV 306  
 RESULT 11  
 Q6DG78  
 ID Q6DG78 PRELIMINARY; PRT; 319 AA.  
 AC Q6DG78;  
 DT 01-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 NCBI\_TaxID=79551;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang U., Hsieh F.,  
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.W., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalon D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman A., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RA Strausberg R.;  
 RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC076473; AAL76473.1; -;  
 KM Hypothetical protein.  
 SQ SEQUENCE 319 AA; 34809 MW; 014DA5E135803B5A CRC64;  
 Query Match 57.9%; Score 961; DB 2; Length 319;  
 Best Local Similarity 63.3%; Pred. No. 5.7e-70;  
 Matches 188; Conservative 44; Mismatches 65; Indels 0; Gaps 0;



Db 253 FFSWFLSKPKEGAQTSYICAVABELQSIGKHFSDCAPAFVAPQGSSEETARKLMVY 309

# RESULT 12

Q9RLRS PRELIMINARY; PRT; 188 AA.  
AC Q9RLRS; 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE UBE-1c1.  
GN Name=9430059D04R1k; Synonyms=Ube-1c;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RC SEQUENCE FROM N.A.  
RA STRAIN=Singapore local strain; TISSUE=Embryo;  
RA Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; MG1:1915364; 9430059D04R1k.  
DR GO; GO:0004022; F:alcohol dehydrogenase activity; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR InterPro; IPR002198; ADH\_short.  
DR InterPro; IPR002424; Insect adh\_fam.  
DR Pfam; PF00106; adh\_short; 1.  
DR PRINTS; PRO1167; INSAHFAMILY.  
SQ SEQUENCE 188 AA; 21263 MW; BA54CC30F43098E7 CRC64;

Query Match 50.2%; Score 832; DB 2; Length 188;  
Best Local Similarity 87.0%; Pred. No. 1e-59;  
Matches 154; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 132 MCPSPKADGEMHIGVNHGHFLTLLEKESAPSRIVNVS;LAHILGRIHFENLQ 191  
Db 1 MCPSPKADGEMHIGVNHGHFLTLLEKESAPSRIVNVS;LAHILGRIHFENLQ 60  
QY 192 GEFKNYAGLAYCHSKLANILFTQELARRLKSGVTVTVHPTVQSELVHSSFMKMMW 251  
Db 61 GEFKNYAGLAYCHSKLANILFTQELARRLKSGVTVTVHPTVQSELVHSSFMKMMW 120  
QY 252 LFSFPIKTPQOGAQTALHCAITGELISGNHFDCHVAVNSAQAENETIARRLMVY 308  
Db 121 LFSFPIKTPQOGAQTALHCAITGELISGNHFDCHVAVNSAQAENETIARRLMVY 177

# RESULT 13

Q6DC71 PRELIMINARY; PRT; 291 AA.  
AC Q6DC71; 01-OCT-2004 (TREMBlrel. 28, Created)  
DT 01-OCT-2004 (TREMBlrel. 28, Last sequence update)  
DE 01-OCT-2004 (TREMBlrel. 28, Last annotation update)  
DE Hypothetical protein.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Singapore local strain; TISSUE=Embryo;  
RA PubMed:12477932;  
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Iqbaljano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Guarnate P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
RA Krzyzinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Singapore local strain; TISSUE=Embryo;  
RA Strausberg R.;  
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC078208; AAH78208.1; --  
KW Hypothetical protein.  
SQ SEQUENCE 291 AA; 32141 MW; B7630F257D7C4386 CRC64;

Query Match 48.2%; Score 799.5; DB 2; Length 291;  
Best Local Similarity 54.7%; Pred. No. 7.5e-57;  
Matches 151; Conservative 54; Mismatches 70; Indels 1; Gaps 1;

QY 34 TSTVOLPKRVVVTGANTGIGETAKELAQGARVYLAQDVEKELVAKETQTTGNOQ 93  
Db 6 TSAARLDGKTVLITGANTGIGETAKELAQGARVYLAQDVEKELVAKETQTTGNOQ 65  
QY 94 VLVRLDLSPTSTISAMAKGFAEKHLHWNNQVWCPKSPKADGEMHIGVNHGH 153  
Db 66 VTISSLDSDSKSISGFAEKHLHWNNQVWCPKSPKADGEMHIGVNHGH 125  
QY 154 FLTLHLLEKESAPSRIVNVS;LAHILGRIHFENLQGEKFNAGLAYCHSKLANILFT 213  
Db 126 FLTLHLLEKESAPSRIVNVS;LAHILGRIHFENLQGEKFNAGLAYCHSKLANILFT 185  
QY 214 QELARRLKSGVTVTVHPTVQSELVHSSFMKMMWLSFPIKTPQOGAQTALHCAI 272  
Db 186 QELARRLKSGVTVTVHPTVQSELVHSSFMKMMWLSFPIKTPQOGAQTALHCAI 245  
QY 273 TEGLEISGNHFDCHVAVNSAQAENETIARRLMVY 308  
Db 246 DPALOTESGKYSDCAPAKAAKAAADDEVAQRLMEL 261

# RESULT 14

Q6NY49 PRELIMINARY; PRT; 331 AA.  
AC Q6NY49; 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DE 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE LOC407663 protein (Fragment).  
GN Name=LOC407663;  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA PubMed:12477932;  
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Ditchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Maira M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Kidney;  
 RA Strausberg R.;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases  
 (SDR) family.  
 DR EMBL: BC066739; AAH6739.1; -;  
 DR InterPro: IPR002198; ADH\_short.  
 DR InterPro: IPR002347; Adh\_short\_C2.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PRINTS: PR00081; GDRDH.  
 DR PRINTS: PR00080; SDRFAMILY.  
 KM Oxidoreductase.  
 FT NON TER 1  
 SQ SEQUENCE 331 AA; 37085 MW; 2AAAE0B547ECC4F CRC64;  
 Query Match 46.0%; Score 762.5; DB 2; Length 331;  
 Best Local Similarity 51.7%; Pred. No. 9.1e-54;  
 Matches 149; Conservative 57; Mismatches 79; Indels 3; Gaps 1;  
 QY 24 IRKMLSSGVCSTVQLPGKVVVVGANTGIGKETAELQARGAVYLACRDVEKGEIVAK 83  
 DB 36 LRRLAGVGRSKARLNGKTVLLTGNTGIGKETAADMAKRGARVILACRDVSRKAAB 95  
 QY 84 EIOTTGNOQVLRKLDLSDTKSIRAMAKGFKAEEKHLHWVNNAGVMMCPYKTDGFE 143  
 DB 96 EIRKSGNENVTVMKDLASLGVRDLVKVQSQEQRDLILINNAGVMCPKHTDEGFE 155  
 QY 144 MHGVNHLGHFLTHLLLEKLSAPSRIVNVSLAHHLGRHFHNLQGEKFNAGLAVC 203  
 DB 156 MGIQVNLGHFLTLNLLDMLKKSAPSRIVNVASVHERGKINFNDIMDKDIDPYQSY 215  
 QY 204 HSKLANILFTQELARLKSGVTTYSVHPGTVOSELVRSSEFMKMMWL--FSFPIKTP 260  
 DB 216 RSKLANVLFRELAIKLRDGTGVTYALHPGVRITELGRHVFSLMKLLILPFYFFKNP 275  
 QY 261 QCGAOTRLHCAITGEGLEILSGNHSFSDCHVAMVSAQAENETIARLMDV 308  
 DB 276 WQGAQTITTCVAVDESLSHSGLYSDCAPKETAPQGRDVAARLMDL 323  
 RESULT 15  
 AAH66739 PRELIMINARY; PRT; 331 AA.  
 AC AAH66739;  
 DT 24-MAY-2004 (TrEMBLrel. 27, Created)  
 DT 24-MAY-2004 (TrEMBLrel. 27, Last sequence update)  
 DE LOC407663 protein (Fragment).  
 GN LOC407663.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_Taxid=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Kidney;  
 KM MEDLINE=2386257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Heien F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uscin T.B., Toshtlycki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Maira M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Kidney;  
 RA Strausberg R.;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC066739; AAH6739.1; -;  
 KM Oxidoreductase.  
 FT NON TER 1  
 SQ SEQUENCE 331 AA; 37085 MW; 2AAAE0B547ECC4F CRC64;

Query Match 46.0%; Score 762.5; DB 2; Length 331;  
 Best Local Similarity 51.7%; Pred. No. 9.1e-54;  
 Matches 149; Conservative 57; Mismatches 79; Indels 3; Gaps 1;

QY 24 IRKMLSSGVCSTVQLPGKVVVVGANTGIGKETAELQARGAVYLACRDVEKGEIVAK 83  
 DB 36 LRRLAGVGRSKARLNGKTVLLTGNTGIGKETAADMAKRGARVILACRDVSRKAAB 95  
 QY 84 EIOTTGNOQVLRKLDLSDTKSIRAMAKGFKAEEKHLHWVNNAGVMMCPYKTDGFE 143  
 DB 96 EIRKSGNENVTVMKDLASLGVRDLVKVQSQEQRDLILINNAGVMCPKHTDEGFE 155  
 QY 144 MHGVNHLGHFLTHLLLEKLSAPSRIVNVSLAHHLGRHFHNLQGEKFNAGLAVC 203  
 DB 156 MGIQVNLGHFLTLNLLDMLKKSAPSRIVNVASVHERGKINFNDIMDKDIDPYQSY 215  
 QY 204 HSKLANILFTQELARLKSGVTTYSVHPGTVOSELVRSSEFMKMMWL--FSFPIKTP 260  
 DB 216 RSKLANVLFRELAIKLRDGTGVTYALHPGVRITELGRHVFSLMKLLILPFYFFKNP 275  
 QY 261 QCGAOTRLHCAITGEGLEILSGNHSFSDCHVAMVSAQAENETIARLMDV 308  
 DB 276 WQGAQTITTCVAVDESLSHSGLYSDCAPKETAPQGRDVAARLMDL 323

Search completed: October 13, 2004, 18:13:07  
 Job time : 77 secs